Chynna Krouser attends a geology class at Hunter College in New York City, where she is a junior with a 3.0 grade point average. Krouser came to Hunter through a pilot project that expanded the admissions criteria for students graduating from performance-based high schools like hers. Krouser submitted an advanced algebra project she completed in her sophomore year at Eastside Community High School in New York City.

ASSESSMENT

EDITOR’S NOTE
States and schools are implementing new ways to measure student learning in lieu of traditional tests. In this Spotlight, learn how educators are effectively assessing group work, how digital games are reducing student stress on formative tests, and consider the merits of grading practices.

CONTENTS

2 Performance Assessment: 4 Best Practices
4 How Digital Games Take The Stress Out of Formative Tests
6 Colleges Crack Open the Admissions Door To Consider Students’ Skills
7 How to Assess Group Projects: It’s About Content and Teamwork
9 Let’s Give Up the Search for ‘Best Practices’ in Grading
10 Five Non-Negotiables in Assessment for Learning
11 Go Rogue, But Be Prepared for a Fight When it Comes to Giving Up Grades

Commentary
Assessment / edweek.org
Published on February 5, 2019, in Education Week’s Special Report: Projects, Portfolios, and Performance Assessments

Performance Assessment: 4 Best Practices
By Stephen Sawchuk

Let’s get this out of the way first: Performance assessment—the idea of measuring what students can do, not merely what they know—is not a new idea in K-12 education.

Teachers have been told to engage students in projects at least since the days of John Dewey, and probably long before that. (The famous Socratic method, after all, requires students to advance and sustain their positions in an argument, not repeat back knowledge.)

Nevertheless, performance assessment has a bit of a riddled history in the United States. In the 1990s—the last major period of experimentation—it was tried at scale and then abandoned in Kentucky, Maryland, and Vermont.

The challenges begin with a definitional problem: Does an essay test count as a performance assessment? What about a short response on an otherwise multiple-choice test? Experts disagree, and such quibbles have fueled confusion over how to measure “authentic” student performance.

The landscape of performance assessment remains hard to parse. Although most testing experts agree that it’s trending again, it’s unclear how widespread performance assessment is. Just ask Jenny Poon, a fellow at the Center for Innovation in Education, which advises four states on the development and use of the tests. She’s worked to create a continuously updated, comprehensive map of state action.

The problem is that, even in states that have policies supporting performance testing on paper, districts vary greatly in how rigorously they implement those ideas.

Still, Poon said a few trends help to raise interest in measuring student performance in richer ways than multiple-choice questions. About 20 states now use the Next Generation Science Standards, she points out, which specifically require students to engage in scientific practices, such as generating hypotheses and recording data from experiments.

A second thrust is states’ interest in better gauging what high school graduates know and can do, as evidenced by the spread of state and local adoptions of diploma seals and capstone projects—presumably a firmer indication of student ability than credit hours.

Experts also know more about performance assessment after years of experimentation. So as you read the report, keep in mind some of the four big lessons they’ve offered up, which are distilled here for you.

1. Decide on goals first.

First and foremost, the experts say: Know why you want the assessment and what benefits you expect to achieve by investing in it.

“There’s no point in teaching someone to write an article for a newspaper and giving them a multiple-choice test to see if they’re able to do that,” said Scott Marion, the executive director of the Center for Assessment, which advises states on testing. “Performance assessment is made for those situations. But if you’re filling in grammar rules, then maybe multiple choice is fine.”

A related issue concerns how the results will be used. Performance assessments are generally more difficult to standardize and less likely to produce comparable results for individual students. That’s probably OK if the test is being used mainly to supplement curriculum or for classroom grading. But it’s a bigger problem if you want to use it for making decisions about whether a student should graduate from high school or for school ratings.

One well-known mishap occurred in Vermont in the early 1990s, when the state’s portfolio-assessment program rolled out. The program used teachers to score collections of students’ best math and writing work. Early results showed that the degree of agreement among teachers’ scores, known as rater reliability, was initially fairly low. In retrospect, RAND Corp. researcher Brian Stecher, who helped evaluate the program back then, wonders whether leaders there got the focus wrong.

“I think what was really beneficial in Vermont was that this broadened to some extent how teachers were teaching mathematics, instead of a reductive ‘I do, we do, you do,’” Stecher said, referring to a common teaching method taught during teacher education.
preparation. “That seems like a good thing to me and valuable in its own right—and might have been a better use of this unstructured portfolio than trying to have it be the basis for a standardized judgment.”

2. Keep costs in mind.

Coming up with good performance tasks can be expensive as well as time-consuming. In short, it’s hard to do performance on the fly or on the cheap. That’s especially the case if what’s valued is the comparability and reliability of scores, which requires creating and field-testing many tasks.

“When you open up assessments to getting students a wide range of response possibilities in terms of format, length, and activities, then it just becomes very hard to manage the time, and materials, and scheduling. It becomes hard to incorporate it into a structured system of assessments, and it also becomes more expensive,” Stecher said.

That’s one reason so few states have done so at scale under federal annual testing requirements. New Hampshire, the sole exception for now, is using some traditional exams in the years it doesn’t administer its locally developed performance measures.

Finally, even if a performance exam is only used locally or for classroom purposes, teachers must invest time and energy to familiarize themselves with its scoring frameworks to make sure they’re grading fairly. Many districts with expertise in performance assessment, in fact, use blind scoring or double reviews of student work—and all that takes time.

And while teachers are generally more knowledgeable about scoring frameworks, or rubrics as they’re called in the field, than they were 20 years ago, there’s still often an expertise gap for teachers who are used to fill-in-the-blanks and true/false questions, said Steve Ferrara, who oversaw Maryland’s now-defunct performance-assessment program in the 1990s. (He’s now a senior adviser at Measured Progress, a testing company.)

3. Prioritize teaching and learning—not just testing.

Performance assessment in education should be part and parcel of reforms to teaching and learning.

Much of the criticism of multiple-choice tests is that they encourage teachers to focus on low-level, easily measured skills. The inverse should be true, too: Give students rich assessment tasks worth teaching to and help support educators to redesign their instruction to boost development of skills like analysis and inference.

In fact, studies from the 1990s on the Maryland State Performance Assessment Program found that under it, teachers had higher expectations for the learning of their students, and principals had higher expectations on what they expected teachers to do. Schools with a high degree of curriculum alignment to the tests showed the most improvement, Ferrara said.

In other words, performance assessment truly requires system change.

“If you don’t include at least parallel reforms in teaching and learning, an assessment isn’t enough,” Marion warns. “You have to improve the meaningfulness of the content, instructional quality, and improve student engagement, too. If you’re not doing those three things, then you’re just rearranging the deck chairs.”

There are also technical reasons why the mirroring of testing and instruction is desirable: Performance assessment hinges on students having had enough exposure to the content and skills needed to complete the task. Otherwise, the assessment might measure generic problem-solving intelligence, rather than how well students grasp and apply what they’ve learned, noted Sean P. “Jack” Buckley, the head of the U.S. Department of Education’s statistical wing from 2010 to 2013, during which he oversaw the development of the agency’s first performance tasks for exams administered as part of “the nation’s report card.”

“This was always something we worried about,” he said. “It is way easier to make a hard test that smart people can do well on than one that shows growth tied to teaching and learning.”

4. Plan for scaling up the exams—and communicating the results.

Parents and teachers can be a performance assessment’s biggest boosters or its toughest foes, which means it’s key to keep them apprised of the assessment program and the logic behind it as it’s piloted, rolled out, and scored.

Teachers, the experts say, should especially be intimately involved in test design and communications.

“It takes time to build the capacity to build quality assessments; it’s almost an apprenticeship approach,” said Paul Leather, who helped get New Hampshire’s performance-assessment system off the ground and now oversees state and local partnerships for the Center for Innovation and Education, a research and consulting group.

“As we built our common tasks, we selected content teacher-leaders who led development of the content and the common tasks,” he said. “Over time, they start to lead the entire system because assessment literacy has reached such a high level, and we believe that actually has to happen for this kind of system to scale. You essentially create a way in which expertise is not just shared as a product, but something that helps others to gain that expertise over time.”

Even when teachers are involved in task design, they can feel left behind without the right training and supports, Ferrara cautioned. “It took so much effort in the first few years [of MSPAP] to get the program up and running that all the investment went into the assessment program and not into professional development, he said. In fact, he recalled, missing materials and a lack of training in the 1992 assessment administration raised teacher ire and got the test slammed in newspapers as the “MSPAP Mishap.”

Finally, as performance assessments yield more-nuanced information on students’ abilities, there’s a related challenge of communicating those results. For six years, Maine required high schools to prepare students to demonstrate competency in eight subjects to earn a diploma. But the experiment faltered in part because districts struggled to communicate what the new grades, often issued on a 1-to-4 scale, meant—and how they’d affect students’ chances of getting into college, according to news reports on the system. By 2018, the pressure caused state lawmakers to roll back the requirements, giving districts the option to return to traditional diplomas.
Dyslexia and the road to graduation

For a child with dyslexia, the journey from the first day of school to graduation is seldom easy, but having a supportive, collaborative team with the right tools can make all the difference.

**How can we deliver a quality educational experience to a student with dyslexia?**

The educational journey for a student with dyslexia can be a tough one. Understanding dyslexia itself, learning to manage daily struggles, and showing strengths versus weaknesses are recurring speed bumps in a learner’s road to academic success. This experience can be more difficult when intervention efforts are not coordinated.

Imagine a framework that instills collaboration between you and your colleagues, while keeping the student’s perspective at the center of the conversation.

---

**View the infographic >>**

---

**Watch the video**

---

**The Road of Education for Students with Dyslexia**

Copyright © 2018 Pearson Education. All right reserved. No reproduction without express written permission.
Published on February 5, 2019, in Education Week's Special Report: Projects, Portfolios, and Performance Assessments

How Digital Games Take The Stress Out of Formative Tests

By Alyson Klein

NEWNAN, GA.

Second grader Brooks Rudnik hunches over a screen, guiding a purple, sunglass-wearing character through an imaginary world called “Keenville.” His mission: help save the planet’s animals, in part by showing off his knowledge of phonics. As he answers questions correctly, virtual jelly beans pile up in a corner of his monitor.

Brooks doesn’t realize it, but he’s taking a test.

Brooks and his classmates here at Newnan Crossing Elementary School are early adopters of the state’s newly developed game-based assessment system for 1st and 2nd graders, a group that tends to get antsy about pen-and-paper tests. Georgia hopes the game—which is also known as Keenville—will eventually act as a formative assessment, giving teachers a real-time picture of how well their students understand math and reading skills like grouping, graphing, and reading comprehension.

Keenville, which was only available in many schools at the beginning of this school year, is still a work in progress. The Georgia education department has released only about a third of the games that will eventually be available. And although the games currently spit out some data for teachers on how their students performed, Newnan’s teachers are anxious to get their hands on more-sophisticated student-data “dashboards” that will give more detailed information on student performance on a range of skills and standards.

All Brooks is interested in, though, are those jelly beans, the game’s digital reward for correct answers. Students can exchange them for accessories for their avatar in the game—called a “Keen”—or to buy accessories for the Keen’s house. Possibilities include a purple lava lamp, an electric guitar, or a fluffy couch.

Brooks’ teacher, Brandi Cook, sees a lot of potential in Keenville.

“They love Keenville. They think decorating those houses and their [character] is so much fun and they have no idea that those problems they’re solving are really assessing them at all. That’s amazing, to have them actively engaged in taking a test and them not know it.”

Brooks, she said, told her that he thought there would be “lots of girls around” if he did well in the game and created a snazzy house for his avatar.

‘Something That Would Be Ours’

The Keenville test grew out of recently passed legislation which called for Peach State districts to develop a system of formative assessments for K-5. After getting feedback from district leaders, Georgia’s education department decided to create a game-based test to help districts meet that requirement for the early grades. The state collaborated with the Georgia Center for Assessment at the University of Georgia and FableVision Studios on the project. Districts that opt not to participate in Keenville must use another formative assessment with their 1st and 2nd graders.

So far, about 200 schools are using the games across 76 districts, roughly half the districts in the state. More districts and schools will likely jump on board as more games are released, said Jan Reyes, the director of assessment development for the state education department.

Right now, there are 10 online assessment games that are part of the Keenville suite available to schools. The state’s vision is to eventually have 31 games by next fall, including 16 for English/language arts and 15 for math.

There are plenty of off-the-shelf options for game-based assessments. Yet Georgia chose to build its own in part to avoid costly annual licensing fees, making the game more financially viable over the long term.

“We wanted to do something that would be ours and we would own it and we could sustain it over time,” Reyes said. She hopes the game can also be more easily customizable and responsive to teachers’ needs.

Eventually, Keenville will incorporate games set in make-believe regions that look like different parts of Georgia—
mountains, a coastal area, a big metropolis, and even a swamp modeled on the state’s signature Okefenokee. The idea is for students from all over the state to see their own community reflected in the game, Reyes said.

One Keenville game—based at a carnival—helps students tell time. Another takes place in a cave where students can unlock a “treasure chest” by reading non-fiction passages on subjects like “air” and then answering questions. Another, called “Peaching Cafe,” asks students to figure out how much food the planet’s animals need, in part by using a counting strategy.

“They’re solving real-world problems,” Cook, the teacher, said.

That’s by design, said Scot Osterweil, the research director of Comparative Media Studies at the Massachusetts Institute of Technology, who is consulting with Georgia on the design of Keenville.

“What a good game does, in general, is give you an interesting challenge and give you fairly open-ended ways of solving that challenge,” Osterweil said. “What we try to do with all these standards is come up with what are the interesting ways in which kids interact with these ideas and can you give the kid an interesting way of working through this stuff, not just here’s a math … question and did you get it right or wrong.”

‘All About Engagement’

Interest in game-based assessment is “definitely burgeoning,” said Valerie Shute, an education professor at Florida State University who has done extensive research on formative assessment. But she hasn’t heard of another state attempting a game-based approach at this scale.

She said there are still “tons and tons of unanswered questions” about game-based tests like Keenville, but she’s personally a believer. “It’s all about engagement. Right now, when people take tests, it’s so boring and tedious; it’s a turnoff for a lot of people,” Shute said.

But she said she would encourage Georgia to run checks to make sure the test can gauge students’ skills accurately.

“For any assessment to unfold on such a large scale, we need to ensure that validation studies are conducted to make sure that the assessments are measuring what they are intended to measure,” she said. (The state plans to do validation studies, said Meghan Frick, a spokeswoman for Georgia’s education department.)

To be most useful, the test should give teachers a clear picture of how their students perform on specific skills and an understanding of why they missed certain questions. That would be “something that teachers can grab and go with,” Shute said.

For now, the state is in the early stages of implementation, still figuring out where the kinks in the program are and how to grapple with them. The state began piloting Keenville last spring and added additional schools this fall. Although teachers can already get some information about how many questions their students answered correctly, they don’t yet have a detailed picture of their students’ performance.

That will change later this school year, when Georgia is set to release more-detailed dashboards that will give teachers a better idea of how their students are mastering the math and literacy skills the game measures, Reyes, the state director, said. The state also plans to roll out broader dashboards for school and district leaders.

Eventually, each student will receive both an overall measure of his or her reading comprehension that’s used in part to help teachers find books that will be understandable but challenging for the child, and an overall rating of math content knowledge.

At this point, though, Newnan teachers haven’t gotten much training on how to use Keenville as a formative assessment, in part because the state is waiting for the more-detailed dashboards to become available.

“Admittedly, it is very basic right now because we didn’t want to hold up the release of the games until we had everything that we wanted in the dashboard,” Reyes said. “We didn’t do explicit dashboard training because we didn’t have all the features, so we wanted to wait until we had more to train on.”

Georgia’s districts have broad leeway in incorporating the game into their classroom instruction. For now, Cook and Tonya Copeland, another 2nd grade teacher at Newnan, said they have used Keenville to reinforce skills they teach in class and to adjust instruction a bit.

Down the road, they’d like to assign students a game instead of using a more traditional test to check their mastery of a skill.

“Instead of giving them the benchmark or a milestone test or something, I could be giving them a formative assessment that way, with whatever standard I assign them,” Cook said. “You would know what to reteach and who to put in what group.”

If Keenville is successful, the state may consider expanding the game-based assessments to other subjects and grade levels.

“We see it as a model for where we might want to go with some other things,” said Allison Timberlake, the state’s deputy superintendent for assessment and accountability. “We could eventually think of some really neat things we can do for upper grades and science. A traditional assessment model may not be able to fully measure what they are capable of.”

---

Second grader Jace Willoughby plays the online game Keenville at Newnan Crossing Elementary School. The Newnan school is among dozens in Georgia using the game-based testing system with 1st and 2nd graders. The state plans to develop 31 such games by next fall for teachers to use as formative assessments.
Colleges Crack Open the Admissions Door To Consider Students’ Skills

By Catherine Gewertz

Imagine a high school where students skip standardized end-of-course tests. Instead, to pass a class or graduate, they show off the results of big projects they’ve done, such as analyzing why the United States lost the Vietnam War or how geometric patterns can be used to produce solar energy.

Could this kind of testing—known as performance assessment—make it tougher for these students to get into college?

That’s a big question on the minds of educators in K-12 as more schools are getting interested in using projects and portfolios to produce a fuller picture of student learning.

The trouble is that most college admissions officers already must review tall stacks of applications quickly. Few can carve out more time to read long descriptions of students’ work or watch videos of their presentations.

“Performance-based assessments can generate a wealth of information for colleges about what could make a student successful there,” said David Hawkins, the policy director at the National Association for College Admission Counseling.

“But in admissions, it’s like trying to cram an enormous square peg into a small round hole” because most admissions offices are set up to process applications in a swift, formulaic way, Hawkins said.

So how can college admissions officers get a quick and accurate sense of what students from performance-based schools have accomplished? A few projects around the country are trying to answer that question.

A Software Barrier

One of those initiatives, Reimagining College Access, wants to lower a key barrier to considering performance assessments in students’ admission applications: colleges’ software systems.

Hawkins, who’s working on that project, said most colleges use software systems designed to process students’ grades and test scores, but they can’t accept videos, research papers, and other projects. Reimagining College Access, based at the Learning Policy Institute in California, works to create online platforms that can accept those kinds of student work.

With the resources to spend more time on each student’s application, the most selective private colleges are the ones most likely to be able to examine more complex forms of student work. And many have been doing that kind of “holistic admissions” for years. The Massachusetts Institute of Technology, for instance, allows students to upload “creative portfolios” that capture research, visual and performing arts, and maker projects.

Performance-based assessments can generate a wealth of information for colleges about what could make a student successful there.”

DAVID HAWKINS
DIRECTOR, THE NATIONAL ASSOCIATION FOR COLLEGE ADMISSION COUNSELING

In a pilot project that began in 2015, CUNY campuses began reviewing the performance-based projects and work portfolios of consortium students, and teachers’ descriptions of their work, even if those students’ SAT scores fell just below the minimum cutoffs for each CUNY campus.

Preliminary findings of that pilot are promising. Graduates of consortium high schools who were admitted through the pilot are earning grades at CUNY that are as good or better than the average for admitted New York City students, and they are outdoing them when it comes to earning credits.

CUNY officials declined to discuss the pilot for this story. But Michelle Fine, a CUNY professor who’s analyzing the data from that project, said it has allowed admissions offices to “encounter applicants they wouldn’t have encountered before,” and use students’ performance assessments as good predictors of their college success.

Opening a Door

Chynna Krouser is a case in point. Under CUNY’s traditional rules, she probably wouldn’t have been accepted. Her grades put her in the top 10 percent of her class at New York City’s Eastside Community High, a consortium school, but her SAT scores fell shy of CUNY’s requirements.

But when the pilot opened the door for Krouser to share more of her work—including an advanced-algebra project in her sophomore year that she defended, Ph.D.-style, before a panel of adult judges—she was accepted to several of CUNY’s four-year campuses. Now, she’s a junior at one of them, Hunter College, carrying a 3.0 grade point average and aiming for veterinary school.

“I’m glad they could really see me and what I can do,” Krouser said.

Schools that use projects and performance assessments often carry an outsized burden of explaining themselves to colleges to make sure admissions officers fully understand their students’ applications.
 Jerome Furman, Krouser’s high school counselor, embraces his role as explainer-in-chief. He cultivates relationships with college admissions officers and spends time describing the projects his students undertake and the nontraditional way their mastery is assessed.

“There really can be a gate there as far as getting my kids seen unless I have a relationship with someone” in the admissions office, Furman said.

Changing Transcripts

Part of Furman’s job is making sure colleges have—and understand—Eastside High’s “school profile.” The profile is a standard high school document that’s sent to colleges with students’ applications. It captures the school in a concise page or two, explaining details such as the curriculum, grading system, and demographic makeup of the community.

A project based in New England has designed model profiles to help schools that use performance assessments convey their work clearly to colleges. They’ve also designed model transcripts to reflect the nature of students’ work in performance-based schools.

Both models were created with feedback from colleges, said David Ruff, the executive director of the Great Schools Partnership, which is supporting the project undertaken by the New England Secondary Schools Consortium and the New England Board of Higher Education.

The new model transcript provides more detailed information than ordinary transcripts. It uses a 1-4 grading scale for students’ courses. But it also provides grades for crosscutting skills, like problem-solving, and for mastery of specific standards within each subject. In English, for instance, students’ proficiency is graded separately in reading comprehension, reading interpretation, writing range, writing research, discussion, and presentation.

The New England consortium and higher-education board also have tackled a key concern of parents whose children attend performance-based high schools: If our children present nontraditional transcripts to colleges, will they have a tough time getting in?

The two organizations convened colleges and universities in New England a few years ago and got a resounding answer: Seventy-five public and private institutions signed pledges that students from performance-based schools would not be at a disadvantage in the admissions process.

“What we heard from these schools is that different transcripts weren’t a problem,” Ruff said. “Across the country today, colleges are already seeing transcripts that use A-F grading systems, or 0 to 100, or 1-2-3-4, or systems where there are As, Bs, and Cs, and no credit below a C. And transcripts from schools that use narrative descriptions.

“The key thing is that they need the transcripts explained to them,” Ruff said. “They just want to understand.”

Most schools that use performance assessments still make a crucial conversion for colleges: They use a letter or number-based grading system, Ruff said. Sending narrative descriptions of students’ work is “still a heavy lift” for all but a few, highly selective institutions, he said.

New kinds of transcripts and school profiles can help bridge that gap, he said.

---

By Sarah D. Sparks

Group work is a time-tested strategy in many classrooms, but educators are starting to rethink how to evaluate these projects not just on the content students learn, but the skills they have to work in teams as adults.

Collaborative problem-solving—the ability to work with others on new and complex problems—is one of the most highly sought-after skills by employers. It’s required under both the common-core math and reading standards and the Next Generation Science Standards. But it’s also notoriously tricky to pull off a project that builds both students’ cognitive and social skills as they meet content standards.

“As teachers, we assume students know how to collaborate,” said Diana Lowe, a curriculum director for math and science in the Texarkana, Ark., school district, which adopted project-based learning as part of its own shift to the Common Core State Standards.

In practice, though, Texarkana educators found group projects often foun-dered—not because students didn’t understand the content, but because they couldn’t work together successfully. The district has been working with the Buck Institute for Education, a nonprofit which provides teacher training on implementing project-based learning, to integrate both content and teamwork skills into assessing their group projects.

“Part of what makes it risky to launch collaborative problem-solving projects for a teacher is there’s a lack of research on how to set up contexts for problems,” said Art Graesser, a psychology professor at the University of Memphis in Tennessee who studies collaborative learning and problem-solving. “Emotions can be involved; the kids might fight over who gets to do what and ... have conflict because part of the construct is to try to get different people’s perspectives.”

Breaking Down Skills

In 2017, the Program for International Student Assessment released the first worldwide test of students’ collaborative problem-solving skills. U.S. 15-year-olds scored in the top 15 of the 52 participating countries, but fewer than 10 percent had strong collaboration skills. On average, U.S. students knew how to volunteer information or ask for clarification in a group, but they were less likely to be able to handle complex problems, mediate group conflicts, or evaluate the quality of their teammates’ work.

“Employers are asking us for specific things that kids can do ... to be able to solve problems on the road, to communicate well with each other,” said William Brazier, the professional-learning supervisor for the Loudoun County, Va., district, which launched a districtwide project-based and group-learning initiative...
in 2014. “Previously, the question was, ‘What information do I need to know for a test?’ Now that question is, ‘What work do I have to produce that will actually have an application in the world that makes collaboration much more important?’”

More-typical group projects, such as science labs, don’t necessarily boost students’ collaboration skills. In fact, students who spent the most time doing practical experiments in science class performed 31 points lower on average on PISA’s collaborative problem-solving test than students who rarely did so.

That may be because group projects in which the answer is already known can make it easier for individual students to slack off, according to Graesser.

He recommended teachers instead set up problems in which students with different skill sets must come together to solve a new problem and produce something. “It’s very visible when you create something. Each person has to do their part or else it doesn’t work,” Graesser said.

The Loudoun district now requires its teachers to explicitly teach students collaboration skills as part of introducing project-based learning and has students develop “contracts” laying out roles and agreeing to rules to guide discussions, such as active listening.

“What we think is needed is not simply assessing the teamwork, but it is training, practice, and feedback,” said Stephen Fiore, director of the Cognitive Sciences Laboratory and team cognition researcher at the University of Central Florida. Research has shown that people in teams taught how to distribute expertise and evaluate what they are discussing, such as the pros and cons of the solutions they develop have been found to produce the best work, Fiore said.

For teachers, that means the process of assessing group projects should include “explicitly quizzing [students] on whether they know what their team members are doing,” Fiore said. “When we look at the kinds of teamwork processes, we would break it down into: How well are they sharing information? How well do they recognize the roles the team members are taking on? Are they trying to meet the goals that the team has identified? How well are they addressing any conflict?”

### Building Trust

Teachers in the Texarkana district mapped out both individual benchmarks for content in each project and a “soft-skills rubric,” which they use to monitor students’ communication, creativity, and teamwork over the course of group projects.

“Even though you’re still measuring students’ individual progress toward the content standards, those soft-skill rubrics do give you a way to look at a team assessment that is not based on the student’s content knowledge,” said Rachel Scott, the director of the magnet program for the Texarkana district.

Teachers build up students’ skills in areas like listening, assigning roles, and monitoring each others’ work in short, low-stakes group activities before moving to major group projects, Texarkana’s Lowe said.

“It’s important as a teacher to establish a [collaborative] culture … so that students develop some trust and appreciation for each other’s abilities and skill sets,” Lowe said. “It sets the tone for them being able to do more with their projects and teams.”

Renee Dooly, a 1st grade teacher at Chico Country Day School, a project-based charter school in Chico, Calif., said she builds in time for class reflections after group projects.

“Kids will be very honest,” Dooly said. “This year’s class … they do have a hard time working together, but they are very good about saying, ‘Oh, this went well because we all had a turn to talk,’ or ‘This didn’t go well because so-and-so wouldn’t participate.’”

Emerging technology may also make it easier for teachers to assess students’ collaboration skills, by allowing teachers to track students’ participation in online planning discussions or edits to group projects.

Dooly said group projects allow students who struggle on traditional tests to show their academic strengths while also getting more support from their partners. For example, in an internal study of eight charter schools in Minnesota and Wisconsin, the EdVisions school network found students’ stronger scores on an assessment of collaboration skills were associated with better math and reading performance.

“You’re also going to have some kids who may not be able to perform at the 1st grade [level] work, but they really are putting the most effort into some part of the group project,” Dooly said. “Group work [becomes] a strength area for them.”

### What Is Teamwork?

**PISA Parses Out the Skills**

The Program for International Student Assessment, or PISA, evaluates a dozen different aspects of collaboration for 15-year-olds across the globe. U.S. students have proven more adept at such group problem-solving than the international average, but girls outperformed boys in every country.

Here’s a breakdown of what that involves:

1. Understanding roles to solve the problem
2. Monitoring and repairing the shared understanding
3. Discovering the type of collaborative interaction to solve the problem, along with goals
4. Identifying and describing tasks to be repeated
5. Monitoring results of actions and evaluating success in solving the problem
6. Enacting plans
7. Discovering perspectives and abilities of team members
8. Building a shared representation and negotiating the meaning of the problem
9. Describing roles and team organization
10. Following rules of engagement
11. Communicating with team members about the actions to be/being performed
12. Monitoring, providing feedback, and adapting the team organization and roles
Your role in the roadmap

When a team of general and special educators, reading specialists, educational diagnosticians, speech-language pathologists, and school psychologists work together using an interdisciplinary approach, the students’ academic success greatly improves. When all of these individuals—working together—have the right tools in hand, the possibilities are endless.

The right tools for the journey ahead

Pearson has created a toolkit for the numerous professionals who serve students at risk for dyslexia—tools that help screen, deeply assess symptoms and underlying skills toward confirming a diagnosis, intervene effectively, and monitor performance at regular intervals. These tools are designed to be used as part of an overall process for dyslexia identification and intervention.

Thank you for helping students find a better way forward

Find out how you can use and implement the toolkit, including more about the empirically validated products designed to support the process of dyslexia identification and intervention. With the right tools and an effective process in place, the educational journey instills confidence in students and educators, and leads to greater academic success.
Early screening is key for student success

Introducing aimswebPlus and Shaywitz DyslexiaScreen

Many states now require dyslexia screening for young students. Schools need quick, valid, reliable, and cost-effective tools to find students who may be at risk for dyslexia. Early intervention is crucial for success in later grades.

More Powerful Together

aimsweb®Plus is a leading research-based screening, progress monitoring, and data management solution for foundational K-8 math and reading skills. It was designed by educational experts and provides reliable, predictive data to inform instruction. With aimswebPlus, schools can identify and group at-risk students, monitor and report student progress, and predict high-stakes testing success.

The Shaywitz DyslexiaScreen™ offers an evidence-based, teacher-friendly observational rating scale for grades K–3. Each rating takes less than five minutes to complete per student. It is intended for mass screenings to sort students quickly—those at risk and those not at risk for dyslexia.

aimswebPlus and the Shaywitz DyslexiaScreen are both included on the list of suggested universal screening tools by the International Dyslexia Association (IDA).
Using the tools together

Option 1: Targeted Screening
Screen all K-2 students with the aimswebPlus benchmark early literacy assessments to flag students at one or more risk levels for reading failure. Then follow up by using the Shaywitz DyslexiaScreen for those students identified as at risk.

Option 2: Universal Screening
Administer the aimswebPlus early literacy battery. Then complete the Shaywitz DyslexiaScreen on all students, six to eight weeks after the beginning of the school year.

After the Screening: What’s Next?
Review data sets for direction on next steps for intervention and/or instructional planning for all students, especially those at risk for reading failure. Or, support those students at risk for dyslexia with further assessment tasks. A ready-to-use handout gives parents/caregivers ideas for supporting students at risk outside school.
Let’s Give Up the Search for ‘Best Practices’ in Grading

By Thomas R. Guskey

It’s time that educators recognize a basic truth about reforming the way we grade and report student learning. That truth is this: There are no “best practices” in grading and reporting, and there is no perfect grading and reporting system.

I know that may be shocking to some and probably unsettling to all. It is surely different from what you might have been told in books, blogs, or chats by writers and consultants who claim to know what is truly “best.” And, quite honestly, I wish their claims were true. If they were true, it would make it so much easier for all of us who are struggling to improve grading and reporting systems and trying hard to ensure those systems consistently work in the best interest of students. But, regretfully, it’s not true.

Susan Brookhart and I recently assembled a group of exceptionally talented scholars to analyze and interpret the vast body of research evidence we have on grading and reporting. Their brilliant work is described in a new book published by ASCD and titled, What We Know About Grading. The conclusion these scholars reached after considering the 100-plus years of research on this important topic confirms that some practices are clearly better than others and provide the basis for making substantial improvements.

We know, for example, that it would be better to replace the appallingly unreliable percentage grading scale that includes 101 distinct levels of performance and offers only the illusion of precision with grading scales that include fewer but more reliably discernable categories of performance. We know it is better to base grades on clearly articulated learning criteria rather than on students’ relative ranking among classmates. We know it is better to offer multiple grades that reflect specific product, progress, and process criteria, rather than combining these into a single hodgepodge grade that confounds interpretation. Making these changes will greatly improve the communicative value of any grading and reporting system.

What this extensive body of research evidence does not offer, however, are prescriptions for specific policies and practices that we know will work best for all students under all conditions, which is precisely what “best” implies.

If we implement grading scales with fewer and more reliably discernable categories of performance, for example, we don’t yet know the “best” number of categories to include or the “best” way to label those categories. If we provide multiple grades reflecting different types of learning criteria, we don’t yet know the “best” criteria to use or the “best” types of product, progress, and process criteria to consider. Although we can take many important steps to make grading and reporting much better, we still have a long way to go in determining what is truly “best.”

The major reason we are unable to identify truly “best practices” and describe a perfect grading and reporting system is the powerful influence of context. What the research reveals is that context matters greatly in grading and reporting, just as it does in almost all aspects of teaching and learning.

What works best varies depending on the purpose, the characteristics of the students (i.e., their age, developmental level, learning background, etc.), the culture and other features of the school, and the nature of what is being communicated. These intervening factors snarl effects and complicate implementation.

Nevertheless, not knowing what is truly “best,” shouldn’t deter us from using the extensive research evidence we have to take positive steps in the right direction in order to make grading and reporting much better. As Mark Manson reminds us:

“Growth is an endlessly iterative process. When we learn something new, we don’t go from “wrong” to “right.” Rather, we go from wrong to slightly less wrong. And when we learn something additional, we go from slightly less wrong to slightly less wrong than that, and then even to less wrong than that, and so on. We are always in the process of approaching truth and perfection without actually ever reaching truth or perfection.”

Our goal, therefore, should not be to employ only “best practices” or to implement a perfect grading and reporting system. We don’t have the knowledge or understanding to do that. The nuances of the process and influence of context make it impossible.

Instead, our goal should be to use the knowledge base we have to do things better than we are doing now. We need to see grading and reporting more as a challenge in effective communication rather than simply one of quantifying student achievement. We need to stop using grades to sort and rank students and instead use them to guide students in making improvements in their learning.

Most importantly, we need to help students understand that grades do not define who you are as a learner but where you are in your learning journey—and where is always temporary. These changes won’t yield grading and reporting systems that are perfect. But they will make those systems much better than they are today—and that alone is important and valuable.

Thomas R. Guskey is a senior research scholar at the University of Louisville.
Five Non-Negotiables in Assessment for Learning

By Gary Chapin

Six or seven years ago, when we were formalizing our approach to “Assessment for Learning” at the Center for Collaborative Education (CCE), we called it Quality Performance Assessment (QPA), meaning performance assessment that achieves technical quality (valid, reliable, sufficient, and free of bias). At the time, we tried to capture the essence of our understanding in a conceptual framework (see figure). Student learning, as always, is at the center, embraced by teacher learning (capacity) and leadership and policy support (conditions). Embedded within and between these, there is an iterative process moving from performance-assessment design to data analysis to aligned instruction and repeat. The graphic is fairly simple, but, six years later, it stands up!

Thoughtful consideration reveals that the QPA framework is founded on a set of assumptions. These five non-negotiable qualities are essential to building a successful performance-assessment system:

1. **Radical Student Engagement**: Student learning is the nexus of the framework, and student engagement is the nexus of student learning. Assessments don’t tell us what kids know. Assessments tell us what kids are willing to show us that they know. Lack of engagement and lack of knowledge—unwillingness to demonstrate and inability to demonstrate—look virtually identical. A quality performance assessment provides opportunities for learning that are relevant and real for the kid. The best way to do this is by allowing the student to genuinely collaborate in the design and/or execution of the assessment: radical learner agency. At the very least, ask yourself, will kids be more likely to succeed because of the design and structure of your assessment or in spite of the design and structure?

2. **Alignment**: Traditionally, validity is an expression of basic alignment. Does the assessment assess the thing you want to assess? Does the learning target align with the instruction rubric, etc.? Alignment has come to mean so much more. Does it align with the student? With their context? With their culture(s)? With their gifts? These are the questions that lead to culturally responsive and anti-bias pedagogies, as well as the primacy of the relationship for learning. These are things that, like everything else on this list, have ceased to be optional.

3. **The Design Cycle**: Performance assessments are not static things that you can look at and say, “This is valid. This is not.” As much as validity depends on the components of the assessment, it depends also on context and implementation. Is this assessment valid for a particular purpose and a particular kid? Are there entry points for every kid? Similarly, reliability depends on a constant examination, in teams, of what does proficiency look like? The calibration process happens later in the game than we would like it to—because it requires student work—but it is easily the most important protocol going. We’ve seen untold educators go through a calibration event—examining student work—with eyes wide as the value of performance assessment, and the genius within their students becomes tangibly evident. And then the teachers do it again. And again. Continuous. Forever. Design, implementation, validation, and calibration are iterative and reflective, not linear and closed. Be willing to be surprised within the infinite loop.

4. **Body of Evidence**: Quality performance assessments require a quality performance-assessment system in order to be truly effective. The naive days of saying one assessment “proves” mastery are long gone. Assessments provide evidence of learning and, combined with other assessments within a system, lead to the creation of a body of evidence for any particular student. From those many points of evidence, sufficiency determinations can be made (is there enough evidence?), and next steps can be planned. The generative nature of the design cycle becomes obvious over the course of time—the collection of assessment data being more than the sum. The QPA framework informs at all levels—student, classroom, school, district, state, etc. One level can-
not function well without the other levels.

5. The Nurturing Learning Culture: In the framework, the triangle of Student Learning is embraced by the circles of Teacher Learning and Leadership and Policy Support. Like Walt Whitman described in Song of Myself, these two circles are “large [and] contain multitudes.” They sum up in a few words many interacting systems. Think of the structural, symbolic, political, and personal dimensions that define a culture (h/t Bolman and Deal). How must this be enacted? Time for PLCs to meet. Freedom for PLCs to determine agendas. Care—in the form, perhaps, of food and recognition—for those teachers who step up. It includes learning for teachers (though not necessarily “training”). It includes book groups, common planning times, space to vent, and the recognition that learning happens one conversation at a time. It requires trust, trustworthiness, and vulnerability—or maybe the vulnerability comes first—from all parties.

The formal Assessment for Learning Project is almost 3 years old. CCE’s Quality Performance Assessment has existed in its present form for over six years. We’ve learned again and again that even the best techniques of assessment are not sufficient for a vital assessment for learning culture. For the deep culture change—for the shift in values, ethics, and pedagogy—that is assessment for learning, these five qualities are non-negotiable.

Gary Chapin is the senior associate of Quality Performance Assessment at the Center for Collaborative Education

COMMENTARY
Published on February 1, 2019, in Education Week Teacher's Work in Progress Blog

Go Rogue, But Be Prepared For a Fight When it Comes to Giving Up Grades
By Starr Sackstein

"I'm no evangelist, so I wasn't looking to convert, but I was hoping to build a solid curiosity where perhaps some folks would at least try to loosen their grips on the grading reins and provide students more opportunities to be involved in the process." - Starr Sackstein

worked alone.

It wasn't my choice to work alone, but no one was coming with me. They all thought I was CRAZY! Some said it to my face others only said it behind my back but loud enough for kids to hear and report back.

Of course, the notion of giving up grades inside of a school system that requires them, where parents and students expect them and other teachers wield them powerfully is definitely outside of the norm, but I was okay with all that.

When I made the decision to give up grades, it was because I knew it would benefit all of my students.

And it did, but it made people uncomfortable and I spent a lot of my time explaining myself and trying to help others understand.

I'm no evangelist, so I wasn't looking to convert, but I was hoping to build a solid curiosity where perhaps some folks would at least try to loosen their grips on the grading reins and provide students more opportunities to be involved in the process. Too often schools focus heavily on the product, missing countless opportunities to increase student learning and efficacy by doing the hard work of revising and growing in class.
Giving up grades isn’t just about not labeling learning with numbers or letters, it is about shifting a mindset about what learning is and how it should look in a classroom at all ages. It is about creating nuanced learning experiences that grow essential social-emotional skills while developing academic skills and building deep content knowledge.

Developing a learning environment that incorporates student voice and choice, engages on a multitude of levels and plays to the strengths of individual children takes time, but the rewards are great.

Here’s what schools can do to help those teachers who are bold enough to make big strides to improve student learning with or without the support of others:

- **Take an interest.** When I started to make my move, my leaders didn’t really understand what I was doing and since things went well, generally and my students were successful, they left me alone. That had its benefits and its drawbacks.

- **Provide feedback.** Everyone needs feedback, even your strongest teachers and administrators. Sometimes it is the encouragement and sometimes it can be as easy as a conversation to bounce ideas off of each other. If you don’t know how to help, that is okay, but maybe you can put them in touch with someone who can. Reach out to your PLC and network for them.

- **Provide resources.** Maybe you read something great that could help. This is an opportunity to give insight through current writing and/or videos that could potentially be very helpful. Maybe you know there is going to be a conference soon with someone notable speaking. Allow the teacher to go and get inspired and get connected. Being supported through these tough adjustments is important and if we can acknowledge that maybe we don’t have all of the answers, helping teachers find others who may have answers is extremely helpful. As I navigated the waters, I reached out to people on Twitter, read anything that had something to do with grading and joined larger conversations. Then I started thinking crazy and taking big risks, all the while including my students in those experiences.

- **Acknowledge progress.** When you see a teacher taking risks, tell you them you notice. Be specific about the growth you see in the students and/or the teacher. Write a note or email or share a vox, but make sure to say something.

It is also cool to ask the teacher, how he or she likes to be recognized and then do that. You can also make connections within the building. Particularly in secondary settings, teachers work in silos and are often unaware of what their colleagues are doing, help them know there are others that could provide support even locally. It’s also important to not make examples of them because there is a good chance their colleagues could resent them.

- **Be a supportive ear.** When a teacher goes rogue, they will often feel like they made a wrong choice to do so. It’s hard and there is usually pushback, so when the teacher is having doubts, make sure to be available to talk them down, reminding them how important the work is and how successful they have been already. It’s easy to lose sight of the successes when you are in the middle of a setback. Hold up the mirror, remind them of how far they have come and that it is okay for mistakes to happen.

If we know people who are bold enough to start changing the way the system works, we need to engage in conversations to understand if we don’t already. We need to be willing to have an open mind and a helpful, supportive outlook. Change is hard, I know it is cliche to say, but it is. We need all the help we can get while we are on the front lines trying to make a difference.

*Starr Sackstein is a secondary educator and school leader in New York. She is the author of several books on education, on topics such as going gradeless, peer feedback, and blogging.*
Editor’s Note:
Cutting-edge schools are leveraging new strategies to upgrade their teachers’ skills through strategic training and more collaborative approaches to professional development. In this Spotlight, learn about ways to personalize teacher PD, tap into social media to transform training, and avoid burnout.

CONTENTS

2 Long Beach Sets Course to Personalize PD Options
3 To Tailor PD, D.C. Looks to Groom Teacher Leaders
5 Math-Modeling PD Takes Teachers Beyond the Core
7 Training Gains Toehold for Aides in Special Education
9 Teachers Turn to Twitter for Solutions, Connection
COmmENTary
10 Extinguishing a Burnout: Actionable Ideas to Keep Teachers Engaged in Their Careers
11 Transforming Professional Development Beyond ‘The Mirage’
13 Give Teachers Time to Collaborate

Newly minted teacher leaders from Bancroft Elementary School in Washington go through an exercise during a summer training session on instructional coaching. The summer coaching intensive is a key component of the District of Columbia’s Teacher Leadership Innovation program, designed to provide structural supports for teacher-leader roles in schools.