The Future Is Performance Assessment

Dan French

Feedback from students and teachers shows performance assessment's potential for improving teaching and learning and better preparing all students for college, career, and life.

> [I excelled in] classes at college where there were required presentations or exhibitions, because at Fenway the science fair, or your Junior Review, or your senior projects, all of these required you to stand in front of an audience and talk about what you had learned, to put it into practice in front of a group of people who are assessing you. (George, Fenway High School graduate, quoted in Gagnon 2010, p. 27)

We are at a propitious time in education in the United States. The Every Student Succeeds Act (ESSA) provides a window of opportunity to re-examine what our accountability systems should look like in the future, a future that looks quite different from fifteen years ago, when the No Child Left Behind Act (NCLB) was enacted. At that time, NCLB and standardized testing cast a new spotlight on achievement disparities by

Dan French is the executive director of the Center for Collaborative Education.

group, a significant development that brought rampant opportunity inequities to the fore.

In retrospect, there were far more shortcomings to NCLB than benefits. Despite the focus on group performance, standardized testing has done little to close yawning achievement gaps based on race, income, language, and disability. Too often – particularly in districts with high percentages of low-income students, students of color, and English language learners – schools narrowed the curriculum and focused on test-taking in order to boost test scores and avoid the punitive labels of being a low-performing school (Pedulla et al. 2003; Crocco & Costigan 2007; Darling-Hammond 2007). External test-making companies created standardized tests that were often divorced from the curriculum, leading to hours lost from learning due to test-prep and test-taking while doing little to build teacher capacity to truly assess student learning.

A WINDOW OF OPPORTUNITY: THE CASE FOR PERFORMANCE ASSESSMENTS

Assessments should test what is most important. David Conley (2012) found that, in addition to content knowledge, colleges seek high school graduates who have intentional patterns of thinking, ownership of their learning, and the ability to adapt to unpredictable change. A 2003 poll for the Association for American Colleges and Universities found that more than 75 percent of employers felt that colleges should

> "place more emphasis on helping students develop key learning outcomes, including: critical thinking, complex problem-solving, written and oral communication, and applied knowledge in real-world settings" (Hart Research Associates 2013, p. 1).

Most important, though, is the fundamental premise that public education should prepare students to be contributing members of a democratic society. Eleonora Villegas-Reimers (2002) notes,

> "Citizens must develop democratic abilities and skills, moral values that reflect democratic ideals and principles, motivation to get involved and act, and knowledge of democracy, its principles and practices" (pp. 1–2). She describes the democratic values citizens must learn: "respect and tolerance (both individual and political), responsibility, integrity, self-discipline, justice, freedom, and human rights" (p. 3).

Measuring these outcomes is far beyond the scope of a standardized test. This is where performance assessment enters the picture. The Center for Collaborative Education (CCE) defines high-quality performance assessments as

> "multi-step assignments with clear criteria, expectations and processes that measure how well a student transfers knowledge and applies complex skills to create or refine an original product" (CCE 2017). For example, a task created by a New Hampshire tenth-grade science teacher to assess students' knowledge of cause and effect required students to create a simple machine with a predicted measurable outcome. A proficient response to the task must have a testable hypothesis, a detailed visual representation, and a plan that accounts for all the major principles involved with an investigation to determine the work completed, efficiency, and mechanical advantage of the machine. Similarly, a science task designed to assess fourth-grade students' understanding of the properties of energy requires students to construct a solar cooker that increases the temperature by a certain number of degrees by developing and testing prototypes, and then analyzing and reporting on their data.

Multiple researchers have found that well-constructed performance assessments are better able to measure higher-order thinking skills while accommodating a wider variety of learning

styles than standardized tests (Darling-Hammond & Pecheone 2009; Niemi, Baker & Sylvester 2007; Wood, Darling-Hammond & Neill 2007). While changes may be imminent under the new federal administration, the current ESSA provides new opportunities for performance assessment to assume a larger role in state accountability models. States are now required to use three academic indicators – performance on state tests, English language proficiency, and a third indicator of the state's choice. In addition, section 1204 enables up to seven states to receive approval to create and use local assessments, similar to New Hampshire's PACE initiative.[1]

"[Performance assessment] forced me to collaborate with different people, different writing styles, different thinking styles. And it really prepared you for a lot of things that you'll do later on in life."

—Janelle, Boston Arts Academy graduate

Student voices

Perhaps the best evidence that performance assessments make a difference comes from students themselves. In 2010, CCE researchers interviewed more than ninety former students who had graduated from three Boston pilot schools where performance assessments were a cornerstone, asking the simple question: "How did attending a performance assessment school help or hinder you?" Almost unanimously, graduates reported that performance assessments had helped them better navigate college, career, and life by teaching them how to problem solve, collaborate, and analyze (Gagnon 2010).

> When it came down to writing research papers and any paper academically, I thought that Fenway really did prepare me to write those papers. . . . [Fenway] always talked to you about your PERCS [Perspective, Evidence, Relevance, Connections, Supposition], . . . and so, in my [college] papers, I always went back to that. Whose perspective is this from? What's the relevance? What's the evidence? (Lisa, Fenway High School graduate, quoted in Gagnon 2010, p. 20)

Engaging in curriculum-embedded performance assessments developed students' skills in collaboration and thinking in new ways:

> It forced me to go outside of my comfort zone. It forced me to collaborate with different people, different writing styles, different thinking styles. And it really prepared you for a lot of things that you'll do later on in life and later on in different work situations. (Janelle, Boston Arts Academy graduate, p. 26)

Performance assessments enabled teachers to better differentiate instruction based on how individual students learn best:

> You can't learn everything in a book. We had many different types of learning. We'd read a book, but then we'd do a lot of different projects. (Aaron, Fenway High School graduate, p. 1)

1 For more on PACE, see the article by Marion, Vander Els, and Leather in this issue.

Most importantly, performance assessments built students' capacity to learn and think:

> You see what you've done wrong, what you need to do to improve. With RICO [Refine, Invent, Connect, Own], [you] look back at what you've done, understand the mistakes that you made and all the things that you've accomplished and show what you want to do for next year to change for the better. (Damian, Boston Arts Academy graduate, p. 15)

Teachers at the center

Moving toward a school, district, or state accountability system in which performance assessment is the predominant means of determining student proficiency is foremost about returning teachers to the center of assessment systems, which is where they belong. After all, teachers have always created formative and summative assessments for their curriculum. However, within a performance assessment system, teachers must be able to create *valid* curriculum-embedded performance assessments that measure and predict student acquisition of the intended knowledge or skill. Teachers need to score the resulting student work *reliably* to ensure comparability of scoring within and across schools. Doing so ensures that the tasks actually measure student performance on the intended standards and that teachers have a shared understanding of what constitutes proficient student work. Teacher-driven performance assessments, then, become a growth opportunity for teachers to improve their craft through collaboration with other teachers, while also leading to richer learning experiences for students.

Much like anyone gaining proficiency in new understandings and skills, teachers benefit from being introduced to specific tools and professional development opportunities in learning how to build a quality performance assessment system. CCE's Quality Performance Assessment (QPA) program provides teachers with protocols and tools to engage in discourse and accompanying professional development to learn and practice these skills, which include:

- a *performance assessment curriculum planning template* to assist a teacher team to collaboratively create a high-quality curriculum-embedded performance task;
- an *assessment validation checklist* used by an educator team to assess whether a draft task meets the multiple requirements to be considered valid; and
- a *calibration protocol* to assist teacher teams to learn the process of reliably scoring student work.[2]

Such processes lead teachers to reflect and improve upon their work, as a teacher participating in a year-long Qualitative Performance Assessment (QPA) Institute reflected:

> It's important to recognize that through this process I see people going back and revising after the project, versus just walking away and saying, "Oh yeah, next year I should do this." There's that additional step of reflecting on your own teaching.

Another QPA Institute teacher noted the change in teacher collaboration

……..

2 For more on QPA, including resources and tools, see http://cce.org/work/instruction-assessment/quality-performance-assessment/.

through the use of tools such as the calibration protocol,[3] which gives teachers a sense of unity on what constitutes quality work:

> Teams have really bought into the process and started to use the tools to analyze their assessments, really taking student work and reflecting back to the assessment task and the rubric, asking, "Did we truly assess what we meant to assess?" So they went through the [calibration protocol the] first time and realized, "Wait a minute, that's not really what we were wanting to assess, but that's what the students perceived. How do we then get to where we want to be with this assessment?"

As teachers experience the cycle of task creation, validation, administration, and calibration multiple times, they build the capacity to become performance assessment teacher leaders, as another QPA Institute teacher noted: "I have become more purposeful and mindful about what it is that I'm really assessing."

THE FUTURE OF ACCOUNTABILITY SYSTEMS

As more people question the value of standardized testing, the public appetite for a change in the accountability system grows. A 2016 national survey found that "voters consider standardized tests the least important factor in measuring the performance of students," preferring instead to have a multiple-measures data dashboard of student progress (McLaughlin & Associates 2016). In an annual national poll on attitudes toward public schools, 64 percent of respondents stated there was too much emphasis on testing, and testing was ranked dead last on a list of what is most important as a strategy for improving public schools (PDK International 2015).

We also have a more refined idea of how to create performance assessment initiatives at scale, based on lessons of prior, often short-lived efforts. A CCE study reviewed seven different performance assessment scale-up efforts both within and outside the United States, many occurring before NCLB in the late 1980s and early 1990s (Tung & Stazesky 2010). The study identified three critical cornerstones as essential for successful performance assessment scale-up initiatives:

- *robust, sustained professional development* to build teacher capacity to create high-quality, curriculum-embedded performance assessments;
- *technical quality* to ensure that performance tasks are valid and student work is scored reliably; and
- *political leadership and policy support* that enables performance assessment initiatives to be successful and sustaining.

Emerging examples of new performance assessment initiatives take into account past lessons, many of which are discussed in this issue. Several initiatives are taking root at the state level, including: the longest-standing initiative, the New York Performance Standards Consortium; New Hampshire's Performance Assessment for Competency Education; and the Massachusetts Consortium for Innovative Education Assessment.[4] National efforts include the Assessment for Learning Project from the Center for

3 In this process, teachers individually score a piece of student work using a common rubric. They then share their scores for each rubric section, discuss score differences and the reasoning behind scoring decisions, and seek to gain consensus on a uniform set of scores.

4 See articles in this issue by Robinson and Cook; Marion, Vander Els, and Leather; and Kelly and Fearing, respectively.

Innovation in Education and Next Generation Learning Challenges.[5]

The benefits of creating performance assessment accountability systems are clear. As described by Tung and Stazesky (2010):

> Not only did teachers' knowledge and understanding of assessment improve through the use of performance assessments in their classrooms, but . . . this work led to improvements in their instruction and curriculum. . . . In addition, teachers reported improved collegiality in their buildings due to the conversations and sharing encouraged by the use of performance assessments. . . . Finally, most of the scale-up efforts showed improvement in technical quality over time. . . . These initiatives showed that technical quality can improve in the course of a few years, and that once teachers begin to understand and use performance assessments, their enthusiasm for them increases. (p. 42)

While some may claim that there is not yet compelling evidence that performance assessment systems are more effective than standardized tests in improving student learning and closing achievement gaps, consider that fifteen years of NCLB has done little to close achievement gaps (Reardon et al. 2013) and in fact has had the deleterious effects of narrowing curriculum, promoting teaching-to-the-test, and punishing rather than supporting schools. On the other hand, performance assessment systems have demonstrated early evidence of improving both instructional practice and student learning – particularly of higher-order thinking skills, a necessary currency for today's graduates. Transitioning to performance assessments as a measure of student learning has equity at its center, with the goal of enabling a greater diversity of students to demonstrate proficiency in what they know and are able to do.

More research is needed on the impact of performance assessments on student learning. But with an ever-diversifying student enrollment, why wouldn't we go down the path of promise rather than continue to use a system that suppresses creative learning and perpetuates wide gaps in achievement by group?

5 See article in this issue by Sang and Worchel.

REFERENCES

Center for Collaborative Education. 2017. "Quality Performance Assessment," Center for Collaborative Education website, https://goo.gl/CPt2f9.

Conley, D. 2012. *A Complete Definition of College and Career Readiness.* Eugene, OR: Educational Policy Improvement Center, http://www.epiconline.org/ccr-definition/.

Crocco, M. S., and A. T. Costigan. 2007. "The Narrowing of Curriculum and Pedagogy in the Age of Accountability: Urban Educators Speak Out," *Urban Education* 42, no. 6:512–35.

Darling-Hammond, L. 2007. "Race, Inequality and Educational Accountability: The Irony of 'No Child Left Behind,'" *Race Ethnicity and Education* 10, no. 3:245–60.

Darling-Hammond, L., and R. L. Pecheone. 2009. "Reframing Accountability: Using Performance Assessments to Focus Learning on Higher-Order Skills," in *Meaningful Measurement: The Role of Assessments in Improving High School Education in the Twenty-First Century,* edited by L. M. Pinkus. Washington, DC: Alliance for Excellent Education.

Gagnon, L. 2010. *Ready for the Future: The Role of Performance Assessments in Shaping Graduates' Academic, Professional, and*

Personal Lives. Boston: Center for Collaborative Education.

Hart Research Associates. 2013. *It Takes More than a Major: Employer Priorities for College Learning and Student Success*. Washington, DC: Association of American Colleges and Universities. https://goo.gl/SHCfiW.

Niemi, D., E. L. Baker, and R. M. Sylvester. 2007. "Scaling Up, Scaling Down: Seven Years of Performance Assessment Development in the Nation's Second Largest School District," *Educational Assessment* 12, no. 3–4:195–214.

McLaughlin & Associates. 2016. *National Survey of Voter Views on Standardized Tests and School Closure*. Washington, DC: PublicSchoolOptions.org.

PDK International. 2015. "PDK/Gallup Poll of the Public's Attitudes towards Public Schools," http://pdkpoll.pdkintl.org.

Pedulla, J., L. Abrams, G. Madaus, M. Russell, M. Ramos, and J. Miao. 2003. *Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from a National Survey of Teachers*. Boston, MA: National Board on Educational Testing and Public Policy.

Reardon, S. F., E. H. Greenberg, D. Kalogrides, K. A. Shores, and R. A. Valentino. 2013. *Left Behind? The Effect of No Child Left Behind on Academic Achievement Gaps*. Stanford, CA: Stanford Center for Education Policy Analysis.

Tung, R., and P. Stazesky. 2010. *Including Performance Assessments in Accountability Systems: A Review of Scale-Up Efforts*. Boston, MA: Center for Collaborative Education.

Villegas-Reimers, E. 2002. "Education for Democracy," *ReVista: Harvard Review of Latin America* 2, no. 1:36–38, https://goo.gl/AkPPXE.

Wood, G. H., L. Darling-Hammond, and M. Neill. 2007. "Refocusing Accountability: Using Performance Assessments to Enhance Teaching and Learning for Higher Order Skills." Briefing paper prepared for members of the U.S. Congress (May 16), https://goo.gl/hlLgfm.

A FOUNDATION'S VIEW ON THE IMPORTANCE OF PERFORMANCE ASSESSMENT

Nick C. Donohue

Nick C. Donohue is the president and CEO of the Nellie Mae Education Foundation.

Over the last year, I have witnessed a significant increase in interest, energy, and investment in student-centered learning, which is defined by:

- more *personalized learning experiences* that meet learners where they are and in terms of who they are, not where or whom we wish they were;
- *competency-, mastery-,* or *proficiency-based approaches* that demand firm foundations of learning before moving forward to other learning challenges;
- *anywhere, anytime* learning that honors achievements made in a classroom, online, or in the real world; and
- *strong student agency*, where learners have a real voice in collaboration with their teachers.

In a future world of diverse student-centered educational experiences guided by these principles, it is crucial to continually assess learning, which is why advocates of new approaches to learning must support advances in performance-based assessments and educational accountability. Quality approaches to accountability balance intrinsic motivators for adults and schools such as supporting professional judgment, autonomy, and growth, and extrinsic ones such as student outcomes on tests; failure comes when one dominates over the other. Of course, it is not wise to simply ask educators to affirm their own excellence, but an external measure that makes no sense to practitioners will be rejected as not relevant. If we want to hold teachers and students accountable, then we need to make the evidence on which they are judged more legitimate. The good news is that ESSA includes provisions for furthering these more rational accountability approaches, largely because states such as New Hampshire and others are demonstrating that systems that include locally developed performance assessments are viable, reliable, and valid.

In addition, those of us clamoring for a revolution in learning must work to change the way we treat educators and support more rational approaches to teacher effectiveness. We need to be allies in supporting educators just as we aspire to support learners. In a student-centered world, we need to value how well educators know their students. Moving forward, teachers need training and professional development to execute strong formative assessments rooted in developmental theory. Teachers will need to be given time and support to collaborate – not just communicate – with parents and other agents of learning. Advocates for a big change in learning should not only care about better supports for teachers but listen to teachers as they develop and implement important ideas about the future of teaching and learning.