

VUE

2017, NO. 46

ANNENBERG INSTITUTE FOR SCHOOL REFORM • *Voices in Urban Education*

Performance Assessment: Fostering the Learning of Teachers and Students

Performance-Based Assessment:
Meeting the Needs of Diverse
Learners

Rosann Tung

The Future Is Performance
Assessment

Dan French

Case Study: The New York
Performance Standards Consortium

Gareth Robinson and Ann Cook

Reciprocal Accountability for
Transformative Change: New
Hampshire's Performance Assessment
of Competency Education

*Scott F. Marion, Jonathan Vander
Els, and Paul Leather*

A Place-Based Process for
Reimagining Learning in the
Hawaiian Context

*Kau'i Sang and Jessica
Worchel*

Case Study: International
High School at Langley Park

*Frishtah Wassl, Christine
Wilkin, and Maggie Ward*

Teachers, Micro-Credentials,
and the Performance
Assessment Movement

Dan French and Barnett Berry

Case Study: The Capstone Project
at Chelsea High School

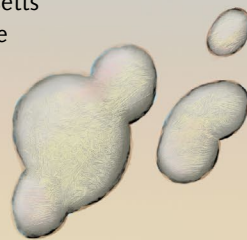
Priti Johari

District, Union, and Community
Collaboration: Massachusetts
Consortium for Innovative
Education Assessment

*Dianne Kelly and
Erik Fearing*

Identity Affirmed,
Agency Engaged:
Culturally Responsive
Performance-Based
Assessment

Ricardo Rosa



ABOUT THE CENTER FOR
COLLABORATIVE EDUCATION

This issue of VUE was planned and produced in partnership with the Center for Collaborative Education (CCE), a nonprofit organization established to promote innovative models of schools and increase justice and opportunity for all learners. The issue was inspired by CCE's work on a new approach to student performance assessment as an alternative to standardized testing. For more information on CCE and resources related to performance assessment, visit <http://cce.org>.

Several CCE staff members collaborated on the planning and production of this issue, including: Gary Chapin, Meaghan Foster, Dan French, Laurie Gagnon, and Christina Kuriacose.

For a glossary of terms related to performance assessment used throughout this issue, see the online edition at <http://vue.annenberginstitute.org/issues/46>.

Performance Assessment: Fostering the
Learning of Teachers and Students
2017, NO. 46

EXECUTIVE EDITOR
MICHAEL GRADY

GUEST EDITOR
ROSANN TUNG

MANAGING EDITOR
Margaret Balch-Gonzalez

ASSOCIATE EDITOR
Sheryl Kaskowitz

PRODUCTION AND DISTRIBUTION
Mary Arkins Decasse

DESIGN AND PRODUCED BY
Office of University Communications

ILLUSTRATOR
Robert Brinkerhoff

Voices in Urban Education (ISSN 1553-541X) is published at Brown University, Providence, Rhode Island. Articles may be reproduced with appropriate credit to the Annenberg Institute. Single copies are \$12.50 each, including postage and handling. A discount is available on bulk orders. Call 401 863-2018 for further information. *VUE* is available online at vue.annenberginstitute.org.

The Annenberg Institute for School Reform was established in 1993 at Brown University. Its mission is to develop, share, and act on knowledge that improves the conditions and outcomes of schooling in America, especially for children who have been underserved by our inequitable systems. For program information, contact:

Annenberg Institute for School Reform
Brown University, Box 1985
Providence, Rhode Island 02912
Tel: 401 863-7990
Web: <http://annenberginstitute.org>

Twitter: @AnnenbergInst

Facebook: <https://www.facebook.com/AnnenbergInstituteForSchoolReform>

© 2017 Brown University, Annenberg
Institute for School Reform

VUE

Performance Assessment: Fostering the Learning of Teachers and Students

Performance-Based Assessment: Meeting the Needs of Diverse Learners

3

ROSANN TUNG

Performance assessment, a personalized and rigorous alternative to standardized testing, provides an opportunity for teachers to build on individual students' strengths and foster more equitable learning outcomes.

The Future Is Performance Assessment

6

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.

Case Study: The New York Performance Standards Consortium

14

GARETH ROBINSON AND ANN COOK

The story of the Institute for Health Professions at Cambria Heights illustrates the positive impact of using performance assessments rather than relying on the state Regents exams.

Reciprocal Accountability for Transformative Change: New Hampshire's Performance Assessment of Competency Education

20

SCOTT F. MARION, JONATHAN VANDER ELS, AND PAUL LEATHER

In New Hampshire, a new performance assessment system focuses on reciprocal accountability and shared leadership among teachers and leaders at the school, district, and state levels.

A Place-Based Process for Reimagining Learning in the Hawaiian Context

26

KAU'I SANG AND JESSICA WORCHEL

Leaders from the Office of Hawaiian Education reflect on their process in developing a culturally responsive assessment framework rooted in Hawai'i's indigenous context, values, and beliefs.

Case Study: International High School at Langley Park

33

FRISHTAH WASSL, CHRISTINE WILKIN, AND MAGGIE WARD

An international school shares their process for developing performance assessments for English language learners.

Teachers, Micro-Credentials, and the Performance Assessment Movement

37

DAN FRENCH AND BARNETT BERRY

Micro-credentials, a new form of personalized professional development for teachers, offer a unique solution to the challenge of training school staff to design and implement performance assessments.

Case Study: The Capstone Project at Chelsea High School	44
PRITI JOHARI	
Capstone projects serve as the centerpiece for an assessment system that values problem-solving and frames assessment as learning opportunities for students and teachers.	
District, Union, and Community Collaboration: Massachusetts Consortium for Innovative Education Assessment	48
DIANNE KELLY AND ERIK FEARING	
Consortium leaders discuss how their model – based on collaboration among districts, teachers unions, and community organizations – aims to change the way school quality is assessed.	
Identity Affirmed, Agency Engaged: Culturally Responsive Performance-Based Assessment	55
RICARDO ROSA	
Performance assessments must be culturally responsive in order to truly serve the needs of students from all backgrounds.	

Performance-Based Assessment: Meeting the Needs of Diverse Learners

ROSANN TUNG

Performance assessment, a personalized and rigorous alternative to standardized testing, provides an opportunity for teachers to build on individual students' strengths and foster more equitable learning outcomes.

“**M**y eyes glazed over at the reading passages, and I had no idea what the multiple choice questions were about. They try to trick you by making all the answers sound right. It was so boring that I didn’t even try to do my best,” said my tenth-grade daughter the evening after taking the PSAT at school.

A recent *Huffington Post* article by a poet whose work was used in the Texas state middle school assessments underscored the inanity of this type of testing. The poet herself wrote that she did not know the “correct” answers to the questions on the test about her motivations for using stanza breaks, similes, capitalization, and imagery in her own poems. “These test questions were just made up, and tragically, incomprehensibly, kids’ futures and the evaluations of their teachers will be based on their ability to guess the so-called correct answer to made-up questions.” She implores all stakeholders, in all caps, to “STOP TAKING THESE TEST RESULTS SERIOUSLY” (Holbrook 2017).

My daughter is a visual and kinesthetic learner in Boston Public Schools. She is creative, hardworking, and inquisitive, but she does not show most effectively what she knows and can do on traditional paper-and-pencil tests. Her current school, Fenway High School, emphasizes project-based learning and uses performance assessments such as papers, skits, presentations, and debates to determine students’ mastery of content. Students have choices in what they produce, so that they are more engaged in the assignment, which is often rooted in the social, cultural, and everyday lives of teens. Examples include a critical gender and race analysis of a popular music video, a propaganda poster on a topic of her choice (body image), and a policy memo on how police departments could reduce incidents of police brutality against Black and Brown people. The culturally

Rosann Tung is the director of Research & Policy at the Annenberg Institute for School Reform at Brown University.

responsive pedagogy elevates critical dialogue, collaboration, visual representation, and inquiry, all of which have been shown to be effective instruction and assessment practices (Piazza, Rao & Protacio 2015).

Prior to Fenway, my daughter attended several traditional schools, in which test preparation and testing were the norm and occupied a great deal of instructional time. Homework included mind-numbing exercises with multiple-choice questions in the form of Massachusetts Comprehensive Assessment System (MCAS) or SAT. Last year, she attended one of Boston's exam schools, a selective public high school where admission is based entirely on a student's grades and entrance exam score. Her principal boasted that parents were happy with the school's assessment practices, which did not need to change to meet the needs of diverse learners, because "our students get high SAT scores." However, SAT scores correlate most strongly with family income and education levels (College Board 2013), not the amount of test prep or the "intelligence" of the test taker!

Fueled by the No Child Left Behind Act (NCLB) and its focus on standardized testing, the U.S. assessment system has been driven by capitalism rather than educational benefit. Annually, the testing industry, which four companies monopolize, is valued at between \$400 and \$700 million. The testing industry drives Americans to spend \$13.1 billion each year on test preparation. Besides the test makers, scorers, and preparation companies, this system is designed to advantage three primary stakeholders: (1) the testing industry's corporate executives, who earn in excess of \$1 million annually; (2) education technology companies, which create online software applications for textbooks, workbooks, curriculum development, formative assessment, and the like; and (3) families, predominantly White, who have the resources to avail themselves of the courses, programs, software, and exposure that lead to higher standardized test scores (Strauss 2015; Alexandra 2016).

My daughter is not alone in her negative experience of traditional assessment. In public schools that are increasingly diverse ethnically, linguistically, and culturally, achievement measurement of the type born of NCLB becomes not only meaningless, but also indefensible: "The acceptance of the reality of diversity is to undermine the possibility for standardized, mass-produced, universally applicable measurement instruments" (Hilliard III 2004).

In this issue of *VUE*, we propose an alternative to standardized testing, whose purpose is to sort and rank students and schools. This alternative, performance assessment, is personalized and rigorous, and improves teaching and learning – thereby benefiting both students and teachers. Against a backdrop of the opportunities provided by the Every Student Succeeds Act (ESSA) and the challenges of a Trump/DeVos education administration seemingly committed to privatizing public education, performance assessment is an opportunity for public schools and districts to better meet the needs of all students and to use more relevant, engaging curriculum and instruction that prepares students for complex problem-solving and collaboration.

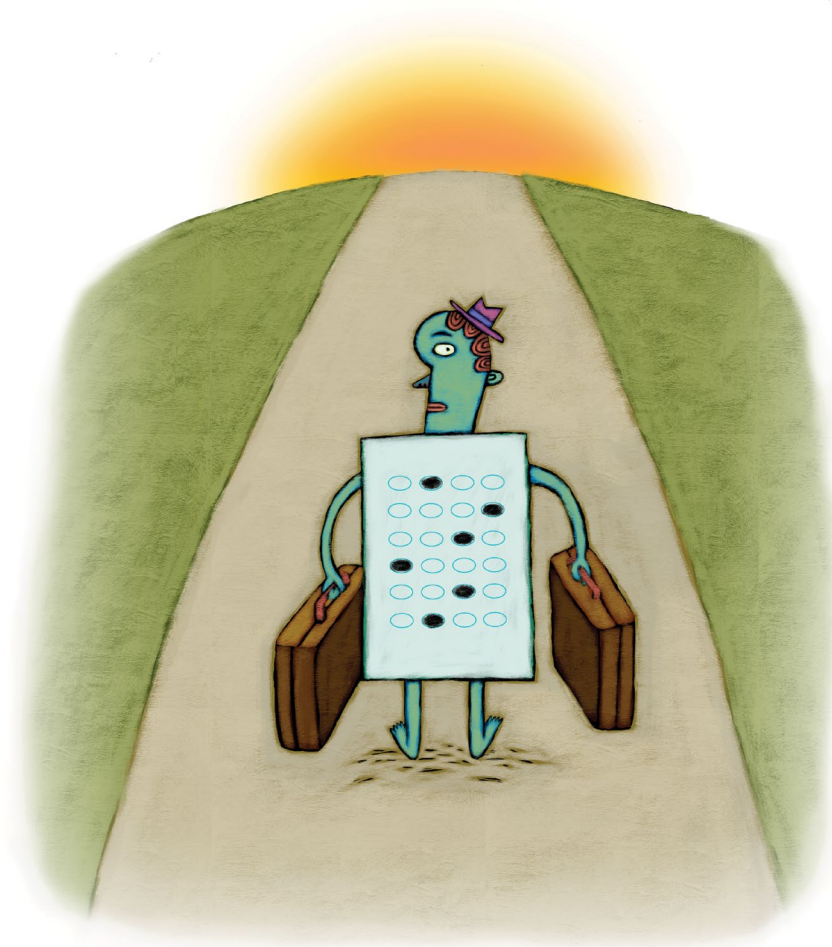
The connection between performance assessment and equity remains a hypothesis. We know that standardized tests exacerbate opportunity gaps. Whether performance assessments reduce opportunity gaps and lead to greater equity depends on how they are implemented and used in instruction. Currently, too little evidence exists that performance assessment closes the "achievement gap" for students who have been historically marginalized. However, given that performance assessments

provide increased learning opportunities and deeper engagement, we expect that students who have been underserved by our inequitable systems will do better with performance assessments than with standardized tests, both to inform instruction and to make decisions regarding promotion and graduation. Some articles in this issue of *VUE* highlight how students like English language learners, Native Americans, students of color who live in poverty, and refugees benefit from performance assessments. Other articles focus on supporting implementation of performance assessments through teacher collaboration; school, district, and state networks; innovative uses of technology; and customized, teacher-led professional development.

My hope is that this compilation of perspectives educates and inspires practitioners, researchers, and advocates to make performance assessment systems the norm rather than the exception – not only for my daughter, but for all students with diverse histories and learning styles and for their teachers, whose dialogue, agency, and learning would be transformed.

REFERENCES

- Alexandra, R. 2016. “The Business of Standardized Testing,” *Huffington Post* (April 24), <https://goo.gl/JvEyNk>.
- College Board. 2013. *College-Bound Seniors: Total Group Profile Report*. New York: College Board. <https://goo.gl/Lz0Hcq>.
- Hilliard III, A. G. (N. B. Amankwatia II). 2004. “Assessment Equity in a Multicultural Society,” New Horizons for Learning website, Johns Hopkins School of Education, <https://goo.gl/T2c6zp>
- Holbrook, S. 2017. “I Can’t Answer These Texas Standardized Test Questions About My Own Poems,” *Huffington Post* (January 5), <https://goo.gl/jzmGxk>.
- Piazza, S. V., S. Rao, and M. S. Protacio. 2015. “Converging Recommendations for Culturally Responsive Literacy Practices: Students with Learning Disabilities, English Language Learners, and Socioculturally Diverse Learners,” *International Journal of Multicultural Education* 17, no. 3:1–20.
- Stecher, B. 2010. *Performance Assessment in an Era of Standards-Based Educational Accountability*. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education. <https://goo.gl/MWKhJp>
- Strauss, V. 2015. “Five Reasons Standardized Testing Isn’t Likely to Let Up,” *Washington Post* (March 11), <https://goo.gl/Dc7bva>.



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

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)



“[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

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

⁵ See article in this issue by Sang and Worchel.

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?

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/SHCfW>.
- 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.

Case Study: The New York Performance Standards Consortium

GARETH ROBINSON AND ANN COOK

The story of the Institute for Health Professions at Cambria Heights illustrates the positive impact of using performance assessments rather than relying on the state Regents exams.

MY PATH AWAY FROM THE TEST (GARETH ROBINSON)

I was born in Port of Spain, Trinidad and Tobago, West Indies. When I was three, my mother, brother, and I immigrated to the United States, joining my grandmother and two uncles in a one-bedroom apartment in Martin Luther King Towers, a public housing development in northwest Washington, D.C. My grandmother and mother worked for a wealthy family whose children attended an independent school, and soon my mother became determined that her sons, too, should attend an independent school.

Ultimately, my older brother was accepted at Sidwell Friends, a Quaker school now known as the school favored by the children of the Washington elite; I joined him when I

entered fifth grade. Before that, I had attended a parochial school and then Adelphi Elementary, a public school in Prince George's County where my mom moved us to pursue a better life. Although I had been accepted to the gifted and talented program at Adelphi, my mother had not given up on her dream of an independent school for me. I remember missing school so I could be tested and visit different independent schools.

I was excited to start Sidwell, because I had seen during my visit that the school had so many things that Adelphi did not. Beyond the physical plant, everyone seemed to know everyone, and everyone I met seemed very interested in me as a person. Reflecting back on it, I believe that this was because the school's community was built on the Quaker core value that an "inner light" exists in all people.

Gareth Robinson is the founding principal of the Institute for Health Professions at Cambria Heights. Ann Cook is the executive director of the New York Performance Standards Consortium.

Two things stand out from my time at Sidwell: first, many of my teachers did not follow the textbooks, and second, we spent a substantial amount of time discussing material. Ms. Reinthaler, to this day my favorite teacher, jumped around the math book in unpredictable ways and was obsessed more with what we were thinking than with the answer we wrote down. We were more likely to go outside and use a cigarette lighter shaped like a parabola or use a ruler sticking out of the board to demonstrate the z -axis than we were to do every problem in the textbook. Ms. Reinthaler's class created a long-lasting impression on me.

Sidwell's classes were interdisciplinary. In English, for example, we spent classes analyzing literature and looking for connections between a particular literary work and related social topics. We looked at what *The Canterbury Tales* had to say about the role of women during Chaucer's time; reading *Native Son* led to a tearful discussion on race relations at the school. Science classes, beginning with biology, included labs and the use of Excel and Word to write lab reports because "that is what scientists did."

Teaching to the tests

When I first entered the classroom as a teacher in the New York City (NYC) public schools, I expected to teach my students the same way I had been taught at Sidwell. That approach didn't work. The world of teaching and learning in the schools where I was assigned was drastically different from the elite world of Sidwell Friends. Not apples and oranges different. Apples and rhubarb different.

During my first two years, I taught at a School Under Registration Review – identified by the New York State Education Department (NYSED) as a school most in need of improvement – where I struggled to create a classroom

where my students would engage in large-group conversation. During an eleventh-grade unit on *The Adventures of Huckleberry Finn*, I would ask my students to categorize and compare what happened to Huck when he was on the shore versus when he was on the river. I was often met with silence, or "Mister, why don't you just tell us the answer so we can go on?" I was also astonished that my kids lacked literacy and writing skills.

Despite my belief in the importance of discussion, feedback from my assistant principal and colleagues required that I change my practice. I was told that "for the sake of the kids," my lessons needed to be connected to the New York Regents exams, statewide standardized tests required for high school graduation. The teacher regarded by my assistant principal as the best English teacher at the school started every class with an exercise taken directly from the exam: providing a "critical lens" quote for students to interpret, agree with or not, and provide two pieces of literature that supported the interpretation. I faced a dilemma: I believed that I needed to teach my students the way I was taught, but I understood that failing to prepare my students for the Regents exams would amount to professional malpractice *and* prevent kids from graduating.

During the next stops of my NYC teaching career – which included a large comprehensive high school with a low graduation rate and a history of violence, a struggling middle school in one of Brooklyn's poorest neighborhoods, and a small high school that selected its students in part based on standardized test scores – my teaching centered around the role of standardized tests. While teaching at the struggling large comprehensive high school, which would eventually be closed after being named one of the most persistently violent in the state, I

was told by colleagues to make sure that I mentioned the Regent’s exam during an observation or the principal might rate the lesson unsatisfactorily. When presenting students with context for a literary text, I created read-aloud passages on which my students were to take notes and answer questions. All my tests and exams were mini versions of the Regents exams and featured the critical lens essay.

While I did my best to avoid teaching to the test, the reality was that I became focused on making sure that the exam would not prevent my students from attending college. When my students graduated, I always counseled them to make sure they visited their new college’s writing center, since I knew that many of them graduated from high school only able to write a critical lens essay.

Creating a school where we would not teach to the test

After spending twelve years teaching English, I seized an opportunity to create and serve as principal of a new high school that would connect students to a possible career, but also focused on classroom discussion. I partnered with the Institute for Student Achievement (ISA) because there was a philosophical connection between my vision for my proposed school and ISA’s emphasis on career and technical education. The collaboration resulted in the Institute for Health Professions at Cambria Heights (IHPCH), a high school that opened in 2013 with 56 ninth-grade students and has grown to 420 students in four grades.

During summer professional development with the founding teachers, I stressed that student inquiry, interdisciplinary connections, and discussion were essential to developing the school I envisioned. Although I had hired teachers with this vision in mind, it proved very difficult to change the

mindset of test preparation among New York teachers. I showed a video of a Living Environment class (the state-required biology course), where a teacher used an excerpt from a science fiction novel to facilitate a discussion. When I asked my teachers how this video could serve as a model for the work that we would do with our students, the Living Environment teacher asked, “How am I supposed to spend several days discussing this book when I have to cover the Living Environment curriculum so the kids have a chance on the Regents exam?”

Similar complaints came from teachers in other departments. “It would be great,” they argued, “to focus on specific historical time periods or core mathematical concepts,” but the need to ensure that our kids passed the exams was the proverbial elephant in the room. The very same dilemma I experienced as a teacher would become the defining challenge in my new role as a school leader.

This situation would change when one of our school’s coaches connected us with the New York Performance Standards Consortium (see sidebar). This introduction led my school’s founding staff to make a decision that would fundamentally change the trajectory of our school’s development.

IMPACT OF IHPCH’S TRANSITION TO THE CONSORTIUM (ANN COOK & GARETH ROBINSON)

Becoming a Consortium school has had a particularly powerful impact on the role teachers play at IHPCH, and faculty ownership of the process is regarded as critically important. To support teachers in the transition, we attended summer workshops held by the Consortium, visited and observed Consortium teacher practice, and participated in the Consortium’s

HISTORY OF THE NEW YORK PERFORMANCE STANDARDS CONSORTIUM (ANN COOK)

The Consortium was created by a waiver in 1995 by the New York State Education Commissioner. The waiver allowed Consortium schools to graduate their students using a system of performance-based assessments (called PBATs or portfolio assessments) in lieu of four of the five Regents exams.

Today, nearly 30,000 students attend the Consortium's thirty-nine public high schools in New York City, Rochester, and Ithaca. Comparative data have demonstrated these students' success, with particularly significant results for four- and six-year high school graduation rates for students of color.¹ NYC Consortium schools serve a higher percentage of African American, Latino, English language learner (ELL), special needs, and low-income students than the city's public schools as a whole, and Consortium students enter high school with lower math and English test results than city-wide averages. Yet these same students graduate at higher rates than the city average, with a four-year graduation rate for ELL students that is 31 percentage points higher than the city average.

Overall, the Consortium schools have a lower dropout rate than the city schools. Statistics from the New York City Department of Education (NYC DOE) show that 83 percent of Consortium students met or exceeded NYC DOE targets for enrollment in college a full eighteen months after graduation, compared with 59 percent of students in the rest of the city.²

A Spencer Foundation-funded study of teachers who moved from test-based schools to Consortium schools found that teachers in the performance-based assessment environment strongly believe they "learn more about their student's academic needs" and are able to "teach more creatively" and teach "more socially just" and "culturally relevant" curriculum. Researchers also reported that teachers in their second year of teaching in performance-based assessment schools felt their students were "more engaged in school" and that using PBATs made "their students more interested in learning" (Hantzopoulos, Rivera-McCuthen & Tyner-Mullings 2016).

For more information on the Consortium, see <http://performanceassessment.org>.

annual conference. The main form of support I gave my teachers was freedom to experiment with both the "what" and the "how" in their teaching. Our emphasis on inquiry-based teaching and learning and discussion-based classrooms resulted in a strong focus on pedagogy and positive rates of teacher retention. Instead of narrowly focusing on anticipated questions on a standardized test, IHPCH teachers plan curricula for students that ensure that interim assessments (pre-PBATs) are aligned to the same skills students will need to complete the more complex PBAT

challenges. For example, one of IHPCH's graduation-level science PBATs is a project where students engineer a catapult using their knowledge of projectile motion and mass. Science teachers collaborated to ensure that in physics classes, students would gain experience contextualizing a design problem, critiquing the process, testing a design prototype, evaluating

-
- 1 See the online version of this article at <http://vue.annenberginstitute.org/issues/46> for a table with these data.
 - 2 See individual school data at <http://schools.nyc.gov/Accountability/tools/review/default.htm>.

the design, and finally, defending their work in an oral presentation.

The transition toward performance-based assessment and away from Regents exams has been both humbling and empowering for IHPCH teachers: humbling because any instructional or curricular problems could not be blamed on the need to prepare students

for the Regents exam; empowering because teachers have created classes that engage students in ways that are not possible when the Regents exam is the summative assessment. Compare, for example, a global history course description offered at a nearby high school with one offered at IHPCH.

NEW YORK STATE REGENTS COURSE	IHPCH COURSE
<p>Global History</p> <p>Semester 3: A survey course roughly covering the years 1500–1920, focusing on the nineteenth century. Topics include the Enlightenment, the French Revolution, World War I, and the rise of totalitarianism. Students will use a combination of cooperative learning, discussions, and critical thinking to better understand how the forces of industrialization, nationalism, and imperialism combined to create the modern world.</p> <p>Semester 4: This course focuses on the twentieth century from World War II to the present. Topics include the Cold War, post-colonial independence movements and contemporary global issues.</p>	<p>Get Up Stand Up, Global Humanities</p> <p>Semesters 3 & 4: Get Up Stand Up is a performance-based assessment class that focuses on revolutions and human rights through a study of the impact of the Enlightenment on the concept of freedom. The course explores how these ideas influenced the political revolutions of the eighteenth and nineteenth centuries. Students complete a case study of the Holocaust as an example of governmental abuse of power and subsequently focus on human rights violations that have occurred since 1945.</p> <p>The course requires that students complete a research paper and make a presentation.</p>

Impact on students and parents

After the school’s first year, which had included students taking Regents exams, the staff announced that the school would be joining the Consortium and making a transition away from the Regents exams. In presenting this decision to parents and students, the staff emphasized that many of the skills staff wished students to develop, practice, and strengthen could be realized through performance-based assessment, including analysis, modeling possible solutions, strategizing,

building evidence-based arguments, oral and written communication, subject-area competence in health care, innovation, creativity, collaboration, revision, and goal setting. Unlike standardized tests, PBATs also allow for instructional coherence and differentiation in the classroom.

The initial reaction of the students to the transition was a mini celebration. They were excited that we had chosen what they thought was an “easier” path to graduation that did not involve the Regents exams. We cautioned

students that this work would be difficult and even included the challenging nature of PBATs in our recruiting talks to parents. Although prospective parents said they recognized similarities between the PBATs and undergraduate or graduate work, they were concerned. They simply could not believe that colleges would accept their children without Regents exam scores. After all, as graduates of New York State high schools, most of our parents had themselves taken Regents exams. In order to better understand the impact of our assessment system on our graduates, we intend to create an email alumni group so that we can track their successes and struggles in college, which will provide us with more concrete evidence for future discussions with parents.

During our school's second full year, we implemented the use of presentations of learning, which evolved into pre-PBAT work. During these presentations, many students asked if we could go back to being a Regents school because "the presentations of learning were asking them to do too much reading, writing, and discussion." There were tears in the hallway and cries of "How can we get all of this work done?" After our physics graduation-level PBAT was given to our first graduating class in January 2016, one student who had failed both the Algebra I and Living Environment Regents exams but passed the physics PBAT said that although PBATs were more work than Regents, they were more interesting and meaningful because the assessment, an engineering project, was more than simply answering questions on paper.

The transformation from a Regents-driven school to one focused on inquiry-based teaching, discussion, in-depth investigation, and oral presentations is certainly a challenge for both students and staff, but I believe we are on the right track.

Utilizing the Consortium's student focused, practitioner-directed system of assessment doesn't immediately transform classrooms into dynamic centers of learning, but, crucially, it allows us to shift our expectations of children from test-takers to active learners, showing that when given the tools and the opportunity, all students can have the opportunity to engage in the level of discourse that I first experienced at Sidwell Friends.

REFERENCE

Hantzopoulos, M., R. Rivera-McCuthen, and A. Tyner-Mullings. 2016. "In Transition: How PBATs Shape Instruction and School Culture." Unpublished draft of study in progress, funded by the Spencer Foundation and Vassar College.

“One student said that although our interim assessments (PBATs) were more work than Regents, they were more interesting and meaningful – more than simply answering questions on paper.”



Reciprocal Accountability for Transformative Change: New Hampshire's Performance Assessment of Competency Education

SCOTT F. MARION, JONATHAN VANDER ELS, AND PAUL LEATHER

In New Hampshire, a new performance assessment system focuses on reciprocal accountability and shared leadership among teachers and leaders at the school, district, and state levels.

For every increment of performance I demand from you, I have an equal responsibility to provide you with the capacity to meet that expectation. Likewise, for every investment you make in my skill and knowledge, I have a reciprocal responsibility to demonstrate some new increment in performance. (Elmore 2002, p. 5)

Scott F. Marion is the executive director of the National Center for the Improvement of Educational Assessment in Dover, New Hampshire. Jonathan Vander Els is the executive director of the New Hampshire Learning Initiative. Paul Leather is deputy commissioner of the New Hampshire Department of Education.

This concept of *reciprocal accountability*, developed by school improvement expert Richard Elmore, is at the core of New Hampshire’s Performance Assessment of Competency Education (PACE), a competency-based educational approach designed to ensure that students have meaningful opportunities to achieve critical knowledge and skills (see Marion & Leather 2015; Rothman & Marion 2016; New Hampshire Department of Education 2016). For PACE, reciprocal accountability means that local educational leaders are involved in designing and implementing the assessment and accountability systems and receive intense technical, policy, and practical support and guidance from the New Hampshire Department of Education (NHDOE) and other experts in the field. PACE attempts to foster organizational learning and change by appealing to the intrinsic motivation of adults to improve their work rather than relying on top-down accountability and compliance strategies.

Beginning in 2012, New Hampshire worked with the Center for Collaborative Education (CCE) to implement performance assessment literacy training, using professional development and capacity building to lay the groundwork for moving forward. In March 2015, the U.S. Department of Education granted permission to New Hampshire and their advisors from the National Center for Improvement of Educational Assessment (Center for Assessment) to pilot PACE, a new assessment and accountability system with significantly greater levels of local design and agency, with an overall goal to facilitate transformational change in performance that best supports the goal of significant improvements in college and career readiness.

As part of this shift in orientation, the state is supporting a competency-based approach to instruction, learning, and assessment within an internally oriented accountability model, in which those being held accountable have responsibility for co-developing the standards,

measures, and bars set for proficiency. Assessment of competency-based learning almost always requires performance-based assessment, and the information learned through this process will continue to inform the design of the accountability system and, hopefully, better inform school improvement (Hargreaves & Braun 2013).

PACE involves multiple lines of work and multiple players. Here, we use three specific perspectives to provide tangible examples of reciprocal accountability in action:

- The first example – of *shared leadership* – is presented by Paul Leather, New Hampshire’s deputy commissioner of education, who as the official leader of the project had to build a structure based on shared decision making among the state, districts, and external partners.
- The second story – of *building local capacity* and expertise – is told by Jonathan Vander Els, the current executive director of the New Hampshire Learning Initiative and former principal of Memorial Elementary School in Sanborn Regional School District, one of the original PACE districts.
- The last example is presented by Scott Marion, executive director of the Center for Assessment and the lead technical advisor to PACE. He discusses the ways in which the evaluation of *technical quality* of the PACE assessment system is based on the reciprocal notion of supporting expertise among local educators while meeting rigorous psychometric requirements.

THE VIEW FROM THE STATE: SHARED LEADERSHIP AND RECIPROCAL ACCOUNTABILITY (PAUL LEATHER)

Under former Commissioner Virginia Barry’s leadership, the NHDOE has long practiced reciprocal or “shared leadership” for the major decisions in our state’s public education. Barry met with the district superintendents and

other educational leadership groups monthly to discuss major issues such as educator effectiveness, educational innovative practices, and the opioid crises. In particular, shared leadership discussions have addressed assessment and accountability for many years, from the adoption of the Smarter Balanced Assessment Consortium¹ in 2014 to the design of state accountability systems since the onset of No Child Left Behind in 2002. It was at just such a discussion, held within the confines of the state's Accountability Task Force in 2014, where the idea for PACE was born.

The task force, made up of superintendents, curriculum supervisors, teachers, and association chapter directors, discussed the idea of moving to a new kind of accountability system more in keeping with competency-based education. Chris Rath, then superintendent of the Concord School District, said in no uncertain terms, "We can't take on something this innovative without you providing us some space to innovate. With the Common Core, Smarter Balanced, and other efforts all being implemented this year [2014-2015], our educators are overburdened as it is." After some discussion, the group agreed with the idea of advancing a pilot to include volunteer districts, where Smarter Balanced would be implemented only once each in elementary, middle, and high school, and a bank of complex performance tasks would be used in grades and subjects where Smarter Balanced was not administered. In this way, the idea of "space to innovate" was integrated into New Hampshire's accountability system.

This model of shared decision making became the operational norm for PACE. A roundtable was created, made up of field representatives from the original

four participating districts, two external partners (Scott Marion of the Center for Assessment and Dan French of CCE), and NHDOE staff (Deputy Commissioner Paul Leather and PACE State Director Mariane Gfroerer). Originally, this group met at least monthly to address all of the issues of design, planning, professional development, implementation, reporting, and technical quality. Nothing moved forward without the full consensus of the group.

Now in its third year, the pilot has grown to eight districts and one charter school, and the makeup of the leadership team remains the same, with each district or charter school represented at the table. Meanwhile, consistent with the principles of reciprocal accountability, the field leaders and teachers have taken on more and more of the ongoing work of PACE. Eighteen teacher content leaders now facilitate the construction of new common PACE performance assessment tasks in English language arts, math, and science for grades 3–7 and 9–10.

With the NHDOE's support, a new organization has been constructed: the NH Learning Initiative, which serves as an intermediary entity supporting the work of both the field and the Department. Also, the New Hampshire chapter of the National Education Association is supporting another group of teacher leaders to facilitate PACE implementation with fellow educators within and across districts. All of this work is overseen by the PACE leadership team, which continues to meet monthly. Members demonstrate their shared ownership and commitment to the success of the pilot in many ways, including through presentations at district, state, and national conferences and to state government officials.

RECIPROCAL ACCOUNTABILITY AT THE SCHOOL LEVEL (JONATHAN VANDER ELS)

When I served as a principal in one of the original implementing PACE schools, reciprocal accountability was at the core

.....

1 Smarter Balanced and the Partnership for Assessment of Readiness for College and Careers (PARCC) are assessment systems that were developed through collaborations between groups of states and educators in response to new, more rigorous Common Core academic standards adopted by most states in 2010 and 2011. See <http://www.smarterbalanced.org/> and <http://www.parcconline.org/>.

of our vision ensuring that all students achieve at high levels. I and my teachers subscribed to a shared leadership model in which we were together responsible for the success of our students, and we needed to work collaboratively to truly maximize the strength of the whole school.

In order for PACE to be effective, the capacity of all educators in each of the implementing schools must be developed to the fullest extent possible. Teachers must possess deep understanding of content, discipline-specific pedagogy, and well-developed assessment literacy to teach and assess a rigorous curriculum using complex performance tasks. Teachers must also be willing and able to work collaboratively in and across schools to develop shared expectations and vision.

We worked hard to develop a culture in which it was safe to innovate. Teachers were used to (and comfortable with) working either individually or within their school-based team. PACE required teachers across schools and districts to function in a professional learning community, through which they learned how to work together most effectively, how to look at student work, understand data, and most importantly, make changes to their instruction to meet the needs of all learners. Our teachers' role was to embrace the uncertainty that comes with stepping out of their comfort zones, committing to working collaboratively with colleagues, and sharing our learning to benefit all.

PACE came along at the right time for our school and our district. We had transitioned to "competency-based learning" a few years earlier, but our teachers really began to develop their assessment literacy by creating, administering, and refining Quality Performance Assessments, a professional development opportunity provided by CCE and initially made available over the summer by the NHDOE. Because we were already engaged in developing high-quality performance assessments, PACE was a logical and timely opportunity to participate in an assessment and

accountability effort that was not based on a single, standardized measure to evaluate students and schools.

Teachers' capacity and professionalism are at the heart of PACE. Relying on teacher leadership and autonomy to be "in charge" of the project has put teachers back into the driver's seat, determining students' competency and utilizing the data from the performance assessments to provide support, intervention, and extension, as appropriate, in a timely manner. For teachers, the essence of reciprocal accountability is a sense of "being heard." As one of our lead PACE teachers explained:

I think PACE has been successful so far because the people working on the initiative believe in the work. The people in charge listen to teacher feedback and are adaptable. We all understand the importance of the work and want it to be successful because it's what is best for kids.

We all have a role to play in the success of PACE, and all clearly understand the need to work with, and for, each other to support our students.

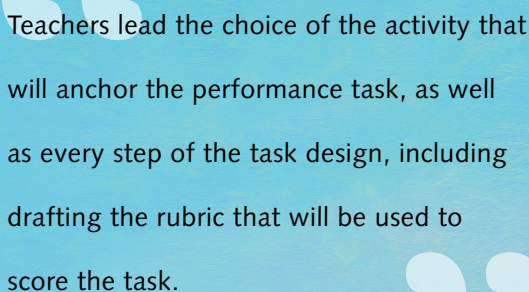
A RECIPROCAL ACCOUNTABILITY APPROACH TO EVALUATING TECHNICAL QUALITY (SCOTT MARION)

PACE has been recognized for its multifaceted approach to the evaluation of technical quality. (See, for example, Lyons & Evans, forthcoming; Rothman & Marion 2016.) In most cases, technical quality evaluations are the purview of highly trained psychometricians like those of us who work at the Center for Assessment. PACE leadership has always had a goal of ensuring that only high-quality assessments were used in participating schools, but we insisted from the beginning of the project that technical quality had to be a participatory sport. In other words, the evaluations of technical quality had to both gauge the quality of the assessments used and to increase the

assessment expertise of participating educators. While there are many aspects of our shared approach to evaluate assessment system quality, we highlight three key components here.

High-quality assessment design

Assessment quality starts with principled and high-quality assessment design. The assessment design templates were drafted by staff at the Center for Assessment, but revised based on feedback and interaction with participating teachers. The Center for Assessment team provides technical support and some oversight to the teacher-led task development teams, but the decisions about which assessments are used in the project are made collaboratively among the teacher leaders, project staff, and the technical consultants. The teachers lead the choice of the activity that will anchor the performance task, as well as every step of the task design, including



Teachers lead the choice of the activity that will anchor the performance task, as well as every step of the task design, including drafting the rubric that will be used to score the task.

drafting the rubric that will be used to score the task. Teachers suggest ways in which the task or tasks will work best within their instructional programs and together with the technical advisors negotiate among district content experts and the technical advisors to design tasks that can serve both instructional and accountability purposes.

Reliable and accurate scoring

Performance assessments must be scored accurately and consistently in order to support their uses to inform instruction and to serve as accountability measures. Further, a key tenet of PACE is that inferences regarding student achievement must be *comparable* across participating districts and between pilot and non-pilot districts, meaning that given a certain set of student work, a student rated as “proficient” in one district would be rated similarly by educators in a different district.

Ensuring scoring quality and comparability starts at the school and district levels, where participating PACE schools engage in calibration exercises to develop a shared understanding of student work quality. The PACE calibration protocol was developed and tested collaboratively among my staff, PACE teachers, and PACE district leads. This process was another example where more top-down technical quality approaches had to be negotiated with the practical realities of doing this work with teachers who have many other responsibilities. For example, we would have liked to have larger samples of student work for our calibration work, but that would have been a burden on the teachers, so we negotiated a sample size that is manageable for the teachers but still provides enough data for us to conduct the necessary technical analyses. In addition to the internal calibration work, each district collects data on the degree to which teachers score the performance tasks consistently with other teachers in the district. The Center for Assessment uses these data to compute inter-rater consistency statistics and then reports back to districts so they can use the information to improve their scoring quality.

Comparability of assessment results across participating districts

The key activity in evaluating cross-district comparability involves a massive collaborative effort led by my psychometric staff and involving hundreds of

educators and project leaders with the main event taking place over the course of two days each summer. Anonymized student papers are distributed to randomly arranged teams of teachers to produce “consensus scores.” These consensus scores serve as benchmarks by which local district scoring is evaluated. (Out of more than 400 papers scored, fewer than five each year required a third rater to help the original raters come to consensus.) Ideally, there should be only small differences between the consensus scores and the scores provided by the original teacher. This alignment would indicate a high degree of scoring accuracy. The more immediate concern is to ensure that the average differences between each district’s local scores and the consensus scoring are similar across districts. The extent to which a district deviates from other districts is a measure of leniency or stringency in local scoring (see Queensland Curriculum & Assessment Authority 2014).

We could have chosen to employ a more typical statistically based approach to comparability, but that would have been more top-down and would have done little to build the skills of participating teachers. The approach we designed allows teachers to collaboratively interrogate student work and to have their consensus judgments play a crucial role in the comparability evaluations. Further, this close examination of student work allows teachers to build their assessment literacy and understanding of student learning.

CONCLUSION

An innovative assessment and accountability project like PACE is unique and important for many reasons. The extensive use of performance assessments helps support learning (Shepard 2000) and increases teacher assessment literacy. The focus on high-quality performance tasks is something we have not seen on a large-scale since initiatives in several states in the 1990s. PACE seeks to demonstrate that some of the past technical concerns with the use of

performance assessments for accountability can be satisfactorily addressed (Evans & Lyons 2017). PACE provides a vivid example of reciprocal accountability in action, framing the ways in which PACE operates at all levels – from the NHDOE, to the approaches for evaluating and improving technical quality of performance assessments, to the collaboration among teachers, to the interactions between teachers and students.

REFERENCES

- Elmore, R. 2002. *Bridging the Gap between Standards and Achievement: The Imperative for Professional Development in Education*. Washington, DC: Albert Shanker Institute.
- Evans, C., and S. Lyons. 2017. “Comparability in Innovative Assessment Systems for State Accountability,” *Educational Measurement: Issues and Practice*. 36, no. 1:1–11.
- Hargreaves, A., and H. Braun. 2013. *Data-Driven Improvement and Accountability*. Boulder, CO: National Education Policy Center. <https://goo.gl/gRYXhk>.
- Marion, S., and P. Leather. 2015. “Assessment and Accountability to Support Meaningful Learning,” *Education Policy Analysis Archives* 23, no. 9. <https://goo.gl/bvASfq>.
- New Hampshire Department of Education. 2016. *Moving from Good to Great in New Hampshire: Performance Assessment of Competency Education (PACE)*. Concord: New Hampshire Department of Education. <https://goo.gl/fPU3hP>.
- Queensland Curriculum & Assessment Authority. 2014. *School-Based Assessment: The Queensland System*. Queensland, Australia: The State of Queensland (Queensland Curriculum & Assessment Authority).
- Rothman, R., and S. F. Marion. 2016. “The Next Generation of State Assessment and Accountability,” *Phi Delta Kappan* 97, no. 8:34–7.
- Shepard, L. A. 2000. “The Role of Assessment in a Learning Culture,” *Educational Researcher* 29, no. 7:4–14.



A Place-Based Process for Reimagining Learning in the Hawaiian Context

KAU'I SANG AND JESSICA WORCHEL

Leaders from the Office of Hawaiian Education reflect on their process in developing a culturally responsive assessment framework rooted in Hawai'i's indigenous context, values, and beliefs.

What would an educational system centered on core Hawaiian values look like?

The Office of Hawaiian Education, established by the Hawai'i Department of Education (HIDOE) in 2015, has been exploring this question through a community-based process that differs significantly from typical Western approaches to policymaking. Often, policymakers use a top-down approach

to policy formulation and implementation that focuses on outputs, outcomes, and impact. In contrast, Hawai'i's new student outcomes framework emphasizes community and indigenous values, knowledge, and shared ownership. This values-based approach is embedded in every aspect of the Office of Hawaiian Education's work – from the student outcomes framework, to the implementation process, to the way they speak about their work. Hawai'i's

Kau'i Sang is the director and Jessica Worchel is the Nā Hopena A'o special projects manager for the Office of Hawaiian Education in the Hawai'i Department of Education.

unique emphasis on community, adaptability, and teaching to the whole child contains transferrable lessons for other policy efforts and contexts.

To understand Hawai‘i’s policy landscape, we must first understand its history. The Hawaiian education system has not always reflected the rich diversity of its population, which encompasses a broad range of cultures, languages, races, ethnicities, and belief systems. Since the 1970s, however, a Hawaiian cultural renaissance has increased the influence of Hawaiian values on policymaking (Wilson 1999). In 2012, the Hawaiian Board of Education formed a working group to strengthen Hawaiian values in the public education system. Educators and community members emphasized the importance of building from the strengths of Hawai‘i, leading the working group to develop *Nā Hopena A‘o* (“HĀ”), a framework rooted in Hawai‘i’s indigenous context. HĀ (pronounced “hah”), meaning “to breathe” or “breath” in Hawaiian, supports a holistic learning process in which outcomes are meant to be demonstrated by everyone within the school system – including students, teachers, and administrators. The Hawai‘i Board of Education approved the HĀ outcomes in 2015, and the state is currently engaging in a two-year pilot with the Assessment for Learning Project (ALP) to develop a valid and culturally responsive assessment framework through a process of *mo‘olelo* [generative storytelling] that draws on the insights, experience, and wisdom of students,

.....

1 The Assessment for Learning Project is a multi-year grant program and field-building initiative inviting educators to fundamentally rethink the roles that assessment should play in advancing student learning. For more information on the partnership between the HĪDOE and ALP, see <https://www.assessmentforlearningproject.org/grantees-hawaii>.

educators, families, and community members.¹

CCE’s Christina Kuriacose and Meaghan Foster spoke with Kau‘i Sang, director of the Office of Hawaiian Education, and Jessica Worchel, Nā Hopena A‘o special projects manager, to learn more about their journey guiding the HĀ framework to implementation and the values that have informed their process. In speaking with them, it became clear that they are treating the new policy as an invitation rather than a mandate, allowing schools and communities to choose how and when they incorporate the HĀ framework into their own context. The Office of Hawaiian Education is deliberately not telling schools what a successful end result will look like; instead, the Office trusts that if schools and communities follow an inclusive, values-based process, they will be able to implement the framework successfully in their own contexts.

To start, we would love to hear the story of the development of HĀ from both of you, and how the HĀ framework differs from other student outcomes frameworks.

Kau‘i: The big question we asked was, “What kind of vision, beyond academic achievement, does HĪDOE have for its public school graduates?” With this question guiding us, our task was to ground our learner outcomes in Hawai‘i *the place*. The general learner outcomes that we were implementing were something you could find in Anywhere, USA. They didn’t tell a story about what it meant to be someone who came from Hawai‘i, lived in Hawai‘i, and was touched by Hawai‘i. Of the twelve of us in the working group, only four of us could speak Hawaiian fluently, but the group in general was attracted to statements that were drafted in Hawaiian.

We took those initial ideas and went through a year-long refinement. We held meetings with educators and community members to discuss what an outcomes model should encompass in order to generate a collective vision. The discourse allowed us to honor the input process and gave us the space to learn how we could strengthen the overall system, piece by piece. With input from all stakeholders, we landed on our final draft in November of 2014, with six core concepts (see Figure 1.)

FIGURE 1. HĀ OUTCOMES GRAPHIC



Jessica: At one of our meetings to get input, the state superintendent stands up and asks, “How can we expect our students to have these outcomes if our system isn’t modeling them? And, more personally, how can I expect my staff to do it if I’m not modeling these outcomes as a superintendent?” That really shifted the conversation from just focusing on students to lifting the new set of expectations up to be system-level outcomes.

Kau’i: We looked to the indigenous culture to understand how we might

shift the whole system. There are so many strengths of Hawai’i. It is one of the most environmentally diverse places on the planet, and its people are also incredibly diverse. We are on an island in the middle of the Pacific Ocean, so we recognize that we must depend on one another and *mālama* [take care] of our people and land, because we are all connected and have limited resources.

This perspective influences our HĀ outcomes to focus on the conditions and learning environments that lead to strengthening HĀ, as opposed to on how to measure an individual’s HĀ. They are about *kākou* – collective success versus individual success.

Jessica: HĀ is different from other outcomes frameworks because it is not just focused on students. Before someone can train others, they must look at how the outcomes resonate for themselves. You can’t force HĀ on others, but you embody and model to strengthen HĀ for yourself and others.

What would you like a graduate to know/believe/embody within your system?

Kau’i: Coming from a native Hawaiian family myself, the importance of accountability to the things that you belong to, to the people that you belong to, is sort of a high-level standard – it’s an expectation which we call ‘*Ohana*. ‘*Ohana* is really what we hope for when we see our graduates move out of the K–12 context into their adult lives, – that they hold space for others.

If you take a look at each of the HĀ statements, they’re really aspirational, and they show up differently depending on the context. What we hope for the graduates – and beyond just the graduates, for all of us – is that they will have the ability to exhibit these outcomes in diverse environments.

Jessica: How do we take care of ourselves as a mental entity, physical entity, emotional entity, and a spiritual entity? How do we allow for us and our education system to honor the whole person, and not just the academic person? Hawai‘i is our place. Wherever we are, that land has something to teach us.

You looked to indigenous culture to help inform this systemic shift. Was that a natural progression, or was there a decision point where the team chose that as a key priority?

Kau‘i: What we are finding is that the doors open when we start to lean on the strength of Hawai‘i first. To your question on whether we intentionally moved to ground the work in indigenous education philosophy, I don’t think we did that initially.

The first iteration of the change in the general learner outcomes policy came out as a very soft translation of typical outcomes in English into Hawaiian. When that iteration was submitted, the board chair and the deputy superintendent pushed back and said that the task may not be to start with what was already there but to create something from Hawai‘i first. Anyone who speaks a different language recognizes that there is a much deeper culture represented in that language and it cannot just be translated one to one. By starting with the Hawaiian language, we started from a different perspective, and therefore the outcome was different.

Jessica: I think that’s something that’s special about the Hawaiian context. Starting with *‘ōlelo Hawai‘i* [Hawaiian storytelling] and honoring the values of Hawai‘i created a more collective outcomes model. The outcomes also include *‘ōlelo noē‘au*, or Hawaiian proverbs, to honor the wisdom of our *kūpuna*, or elders and ancestors.

HAWAIIAN LANGUAGE IMMERSION PROGRAM

Efforts to revitalize the Hawaiian language began in the mid-1980s, when a network of private Hawaiian immersion preschools called ‘Aha Pūnana Leo [the language nest] successfully lobbied the state to reverse the colonial-era ban on the language. In 1987, HIDOE began its own network of public Hawaiian language-immersion schools, called Ka Papahana Kaiapuni. Today, fifteen traditional public schools and six charter schools educate some 2,000 of the state’s public school students in Hawaiian. However, challenges have emerged in creating policies that are effective in both immersion and Western schools, particularly around assessment.

Indigenous cultures place high value on ancestral wisdom. Indigenous perspective also values diversity of ideas, so the final outcomes model, while creating a shared framework and language, allows for a multitude of interpretations based on context. Because of the process being so inclusive and the honoring of the *‘ike kūpuna* [ancestral knowledge], it is our *kuleana* [responsibility] to share HĀ.

Can you speak more about the influence the community has had on the development of HĀ?

Jessica: Public education was initially designed to separate students from the community – whether it is from their language, culture, or community “teachers.” Children would arrive at the school and be asked to leave their community at the door. We are now acknowledging that the community is just as important – if not more so – to the education of the child. We want to help students ground themselves in who they are and where they come from, meaning that our teachers and staff must also become fluent in the

HĀ IN ACTION

Moloka'i High School created a Pu'uhonua Pass that allows students who are having a challenging time being in the classroom to take the pass and go out to reflect on their actions. In Hawaiian, *Pu'uhonua* means "place of refuge." The pass includes questions related to HĀ and allows students time to reflect on their conduct instead of relying on detention.

Kahakai Elementary grounded their Positive Behavior Intervention Supports in HĀ. They recently went through an all-staff orientation and are now reconsidering their essential questions and who their stakeholders are to align with the HĀ framework.

Kalihi Waena Elementary is rolling out HĀ to all teaching staff. They held a professional development day at Ho'oulu 'Āina, a local nonprofit that has been stewarding and sustainably developing 100 acres and is dedicated to cultural education and community transformation. The school is working to deepen their ability to have teachers use the community resources and take students out onto the land to learn.

The **Campbell/Kapolei Complex Area** received a Project Lead the Way grant to bring in community support to make the curriculum more culturally relevant and place-based. They are also taking teachers out to engage with the community and learn more about the native Hawaiian culture.

The new Global Youth Leadership course at **Castle High School** weaves together student leadership, Hawaiian leadership, and global leadership. Co-created by multiple partners, the course incorporates the vision and values of the Polynesian Voyaging Society's *Mālama Honua*,² the HĀ outcomes, and leadership concepts and competencies. Semester units focus on themes of home, destination, wayfinding, and *Mālama Honua* [to care for our Island Earth]. Students participate in indigenous and Western leadership practices, experiential learning, community engagement, global studies, and conferences.

culture and place of the community. Therefore, we must build up our community education space and create room in the system to allow for seamless access between schools and communities.

Kau'i: In the 1990s, the Hawaiian language immersion group created a statewide consortium of parents, teachers, school administrators, community organizations, and state Department of Education staff. The group began to talk about some of the issues facing the Hawaiian language immersion program and collectively try to push on the same issues to create change. As they started to lift that voice into the system more and more, I think the current superintendent saw activism as something to be valued. It started to give the system some answers to the "how": How do we integrate community into education decision-making? How do we share the accountability for the work that we're responsible for, so that it's not just one stakeholder group having to hold on to the weight of a system?

Jessica: We recently hosted a HĀ designers convening and invited teams from across the state to come together in order to learn more about HĀ, share their experiences, and plan to host community days in their region. Each team was asked to bring a staff member from a school, a student, and a community member. We are intentionally working to build and strengthen the connections with the local community and to include students so that teachers can lean on community resources. We are also planning a HĀ Summit, which will bring sixteen school-community teams together to share how they are contextualizing HĀ and determine how to strengthen HĀ within and without the HODOE. On the planning group, we have a mix of internal HODOE staff and external community representatives.

.....

2 See <http://www.hokulea.com/worldwide-voyage>.

We also lift up folks who are not typically looked at as experts or given a voice. For instance, I did a presentation on HĀ at Maui High School, and afterwards one of the skills trainers who works with autistic students came up to me and said, “Well, I’m just a lowly skills trainer, but I would like to have a poster of the outcomes.” I looked at her and said, “You are just as important to this community as any other person in this space.” And same with our clerical staff. We also have our secretaries do presentations with us and talk about their own stories in connection with HĀ, so I think there’s another piece about how we give value to every person in this system who is contributing to our kids.

What do you see as the key attributes of a culturally responsive assessment framework in Hawai‘i?

Jessica: I think the two critical components are: (1) you have to value and honor the indigenous perspectives and indigenous ways of knowing and being; and (2) you have to trust. We honor and value the *mo‘olelo* [stories] of all. Through storytelling and conversation, we make meaning and ensure every voice counts. We must ensure that it is not only one story being told. Currently, the assessment framework is still in the design phase, but I think a lot of folks have difficulty when we talk about multiple pathways with assessment because they expect we have federal regulations and state regulations and a very complex law and compliance system.

Kau‘i: In a town like Waipahu, a town with a large population of Filipino students, they can better design content for their context than a Hawaiian language immersion school that has 99 percent Hawaiian students. The outcomes framework – even though it’s

BUILDING AN ASSESSMENT SYSTEM

HIDOE is in the process of creating a HĀ Assessment Framework through a pilot with the Assessment for Learning Project. The first step of the pilot was a listening tour to generate ideas for developing an assessment system grounded in the HĀ framework. The pilot team is currently testing and refining potential tools and processes that have emerged through their *mo‘olelo* [generative storytelling] process. They hope to complete an expedited second round of testing the tools by mid-summer 2017.

The pilot has provided the HIDOE with the time and space to learn from the Hawaiian context and community. They have learned that it is important to shift the emphasis from assessing an individual student’s achievement, to instead assessing the learning environment and the components that enable students to demonstrate HĀ. Unlike an individual accountability model, the model that has emerged is focused on identifying optimal conditions for building HĀ within learning communities. A picture of a HĀ evaluation system is beginning to take shape as the pilot team continues to seek community feedback, iterate, and re-incorporate Hawaiian wisdom and values.

starting off with that indigenous mindset – really is trying to shift the system. In the context of indigenous cultural practice in Hawai‘i, the *ali‘i*, or chief, actually had a group of advisors who would advise him on the best way to treat the community and take care of the community. It wasn’t his individualistic dictator-style of relationship, but it was really around, “How do I make sure the decisions we collectively make create this sense of *lōkahi*, or balance, in our ecosystem?” We are trying to lift up that practice

and put it into the educational context where this idea of multiplicity allows us to create a much more balanced assessment ecosystem.

It seems like a lot of the values you're speaking about run so counter to current assessment practices, which emphasize a single path to demonstrate knowledge and prioritize individual success instead of collective success. How are you currently talking about the ramifications of that shift with educators?

Kau'i: We think that readiness is a huge factor as we introduce conversations, because it is quite a huge shift in thinking. If you take a look at the entire system itself and the 280 schools and the 180,000 students (including our public charter schools), the range is huge in terms of readiness. We have folks who are absolutely ready, and we have folks in schools who are walking in the opposite direction, and there are a bunch of people that fall in between. When we go out and we share the story, we're asking a particular question about the context, and we're trying to design the presentations and the conversations and the work around that context so that they can see themselves in the work.

Jessica: In terms of the rollout of HĀ, we only go to places where we're invited to talk about HĀ. It's really a grassroots approach. Instead of us going out and doing all this big push or branding or messaging and requiring people to participate, we're allowing folks to ask for an orientation or attend a convening. They invite us into their space or they actively choose to be in our space. They then go back and share with others to build buy-in to the idea of shifting and intentionally incorporating the HĀ outcomes. This way, we know there's already a level of readiness.

While there are guiding indicators, we

ask people to develop what HĀ means in their context. HĀ is about empowering people to define the outcomes and indicators associated with those outcomes for themselves, which builds ownership and accountability. We're shifting our perspective at the state office from being this compliance-driven entity, simply mandating changes in school policy, to a support network.

The guiding principle that we always come back to: Is this best for your students? Are you seeing them inspired? Are you seeing them engaged? Are they learning? As an educator, you know when you see that and feel that, and it's not necessarily a test score.

Concluding with what's best for students and student engagement feels really fitting. Anything else you want to add?

Jessica: HĀ is how you address the achievement gap. You actually create a system that creates the conditions for success for all kids instead of trying to cram those kids that aren't currently being served into a mold they don't fit.

How can you create an education system that really lifts up and values what our kids bring? They're so unique and they're so talented. We need to create that space for them, to support them.

For more information, see <http://www.hawaiipublicschools.org/TeachingAndLearning/StudentLearning/HawaiianEducation/Pages/HA.aspx>.

REFERENCE

Wilson, W. H. 1999. "The Sociopolitical Context of Establishing Hawaiian-Medium Education," in *Indigenous Community-based Education*, edited by S. May, pp. 95–108. Philadelphia: Multilingual Matters Ltd.

Case Study: International High School at Langley Park

FRISHTAH WASSL, CHRISTINE WILKIN, AND MAGGIE WARD

An international school shares their process for developing performance assessments for English language learners.

The International High School at Langley Park (IHSLP) opened during the 2015-2016 school year. By the fourth year of operation, the school will be home to 400 English language learners (ELLs) new to the United States. Working in partnership with the Internationals Network for Public Schools, the school is designed around the “HELLO principles” – five core principles guiding the Internationals’ approach to supporting ELLs:

- Heterogeneity and collaboration;
- Experiential and project-based learning;
- Language and content integration;
- Localized autonomy and responsibility; and
- One learning model for all.¹

It takes three to five years to develop oral proficiency in a new language and

four to seven years to develop written academic proficiency. With only four years of high school, our students need every lesson to be purposeful and accessible (Hakuta, Butler & Witt 2000).

STUDENT PERSPECTIVE: PROJECT GOALS AND GRADES (FRISHTAH WASSL)

Our school, IHSLP, is the best school I’ve gone to. Everything in IHSLP is different. In our school, we do projects. This week I got a project from physical education that taught me how to make my fitness plan. I learned things I did not know, such as the SMART goals formula – setting goals that are specific, measurable, attainable, realistic, and timely.

.....

¹ See <http://internationalsnps.org/about-us/internationals-approach/>.

Frishtah Wassl is a first-year student and Christine Wilkin is a founding teacher at International High School at Langley Park in Bladensburg, Maryland. Maggie Ward is a mastery specialist with the Internationals Network for Public Schools.

Whenever I wanted to exercise, I could only keep it up for about six months, and then I would give up. Now I know the goals that I was setting were too high and not realistic. From this project, I learned that my goals must be SMART. The PE teacher interviewed us on our projects, and she made sure we were on the right track. That's why this is the school I love!

To me, it is really important that I know the goal of the project. In my middle school, I did not know why I got a C in my first quarter, and I was not that willing to ask the teacher for a reason. Now I can see the details about my grades and why I get them. Our grades are not just A, B, C. We are evaluated on how we summarize, analyze, gather information, or make responsible decisions, as well as many other skills. I really like this detailed type of grading, and it helps me reflect on the quality of the project that I turn in.

TEACHER PERSPECTIVE: RETHINKING PEDAGOGY AND MASTERY (CHRISTINE WILKIN)

Upon starting work at IHSLP – with a 100 percent ELL population – I thought I was prepared. I had taught in China and in Haiti. I taught art at a 50 percent ELL middle school, where every ELL student took art as an elective because it was thought to be the easiest class for students like them. And I agreed; art was visual. I could demonstrate the work, and ELL students would be successful.

For the first half of the year I used a lot of the strategies that I had learned to teach ELLs. I had word banks, diagrams, demonstrations, visuals, and examples of what I was teaching. But my students were not engaged, and they were not performing. I thought I was giving them everything they needed to be able to do the task, but students were still not completing assignments. What was wrong?

I started to look at how I organized the lessons. The classic “I do, we do, you do” was not working. ELLs did not understand me when I was speaking in front of a class and demonstrating; they had no background knowledge or language to grasp it. When they transitioned into “we do,” the students with higher language proficiency would do most of the task while those with lower language skills copied. When it was time for “you do,” some students could do it while others just checked out.

At IHSLP, I completely changed the way I was teaching and presenting information. I start with “you do,” to build background information. Then “we do,” to try to understand what they just did together. Then I wrap up with “I do,” and we discuss it.

Now students in my class are presented with a unit made up of five modules and a mastery project. At the beginning of the unit, they read a summary of what they will learn, what tasks they will need to complete, what project they will be making. They will see a map of which skills they will be assessed on along the way. Students start to get the rhythm of the modules, and that consistency allows them to become more independent with their learning.

Within the unit, each module now has a purpose.

- Module 1: background knowledge, struggling, questioning, and experimenting
- Module 2: history and cultural connections
- Module 3: gathering information about the content
- Module 4: start planning their creation
- Module 5: plan their project
- Module 6: mastery project

For example, my students recently worked in groups to make a stop-motion animation movie. They watched “Gumby” videos, built background knowledge by writing about how the characters moved, what materials they observed, and how things were transformed. They went on YouTube and found another stop-motion animation video, researched its history, wrote about what they saw. They looked at tutorials and downloaded video editing software to learn about how to make stop-motion animation, then practiced making one with some photographs that I gave them. So now they understood: here’s an example; here’s the story behind it; now I know how to make it. And then they got into groups to form a production company, and each student had a different role: director, animator, photographer, video editor.

They needed to work together to create the project, and they were being assessed on different skills: planning (submitting a storyboard); critical thinking (gathering information, modeling); ability to reflect and revise; and the content skill (aesthetics and criticism). Then students showed their final projects to the class and critiqued each other’s work, but I’ve been assessing them at each step along the way. All of the modules are uploaded online, and the beginning of each module states what competencies or skill I’ll be assessing, with a rubric that is standard across the school for different skills. I never give a traditional test; it’s graded by the project.

Once students became familiar with how to work through my units, they can independently learn the content with me as a guide, and I can better help students along the learning journey.

COACH PERSPECTIVE: SCAFFOLDING PERFORMANCE ASSESSMENTS ACROSS THE SCHOOL (MAGGIE WARD)

Next up for the presentations is Melvin. He stands tall and proud. He starts by shaking hands with each panelist while displaying a cheeky grin on his face, knowing that he is impressing the panel. He takes a deep breath, exhales, and then starts barreling through words. He speaks at a normal Spanish cadence, rattling off words like a fast-talking New Yorker. About 50 percent are clear English words, 20 percent are Spanglified words, and the rest are likely words in English pronounced in a creative way that my ear is incapable of discerning.

Through it all, a story emerges. There was a hospital visit, a broken ankle, and a surgery that included a drill. He then shifts gears and starts to pass around a series of pictures that show his invention: drill with a shield. I finally realize that the word I keep hearing is “virus.” The guard is to protect against virus during surgery.

This is a prime example of both the beauty and limitations of performance assessment for English language learners. The beauty is that Melvin spoke for five minutes in front of a number of adults and some of his peers and communicated a message in English. His language growth is incredible; in a short time, he has moved from producing words to full sentences.

Watching Melvin speak reminds me of a dinner conversation I had with some Spanish-speaking friends a few weeks prior. The conversation turned to politics and my brain short-circuited. I have strong Spanish comprehension, government vocabulary, and knowledge of politics, but I did not showcase my knowledge because my brain was busy translating. I produced simple sentences like “No, that is a bad idea,” and I could use more complex

SCAFFOLDING CYCLE: MODELING UNIT IN TECHNOLOGY CLASS

Stage 1: Building the Field Students start by comparing bridge models from a previous project using language they already know such as “line,” “this part,” “top,” and “bottom.”

Stage 2: Introducing the Discourse Students are given formal vocabulary and stems such as “length,” “inches,” “longer than,” and “shorter than.” Students create a model and description of a classroom object.

Stage 3: Joint Construction Students swap papers and provide feedback to their partner on how to improve the model and description.

Stage 4: Independent Production/Action Students create and improve their own model.

sentences like compare and contrast or thesis-style statements when modeled by someone else during conversation, but I was unable to produce the formal language on my own. Just as my Spanish limited my ability to showcase my political knowledge, Melvin’s English limits his ability to showcase his medical knowledge.

The question that arises from this scenario is: What do we need to shift in our performance assessments and instruction to give us a clearer picture of what is happening in Melvin’s head?

The first piece we can modify is the performance task. There are two elements to this: modifying directions for comprehension and providing scaffolds to help the students create the output. As Christine mentioned, modifying the directions includes text chunking, word banks, pictures, and pre-reading strategies. Scaffolding the output using multiple modalities, sentence frames, and organizational scaffolds is often the missing link. In the Spanish politics example, my primary output scaffold was repurposing

sentence structures used by others. If I had a reference sheet of sentence stems, my ability to communicate would have been transformed.

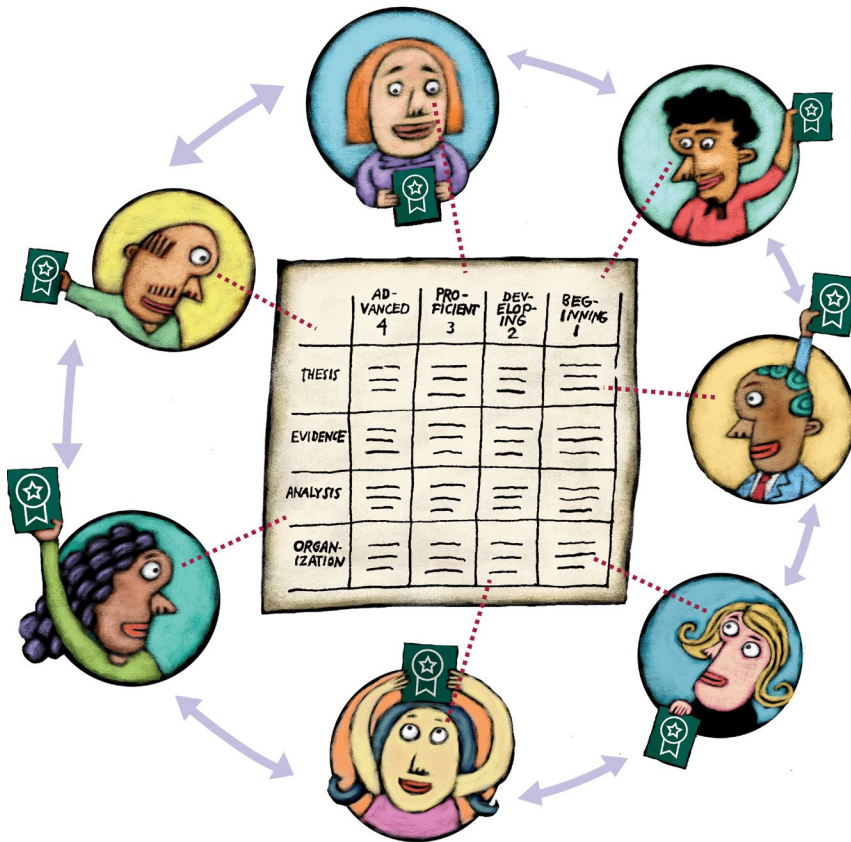
The second way that we lower the language burden of performance assessments is to think in terms of entire units instead of individual assignments, lesson plans, or performance tasks. We use the scaffolding cycle to build entire units to prepare students to engage in performance tasks (see sidebar).

The third manner of addressing this need for scaffolding involves an approach that reaches beyond the individual classroom. It involves creating consistency across classrooms to help students transfer skills from class to class and from one performance assessment to another. While we know that content and vocabulary will change from class to class, language functions needed to describe connections between vocabulary words will not change. The language structures that we need to write essays, give speeches, and engage in debates are the same in science, math, and English.

If we as a staff can leverage those overlapping structures and language functions, we can help our students succeed across classrooms. As a school, we are working toward this goal through weekly interdisciplinary team meetings, school-wide rubrics, peer observations, and walkthroughs. With each conversation, our instruction is more cohesive across classrooms, and our ability to capture the ideas in Melvin’s head becomes stronger and stronger.

REFERENCES

Hakuta, K., Y. G. Butler, and D. Witt. 2000. *How Long Does It Take English Language Learners to Attain Proficiency?* Santa Barbara: University of California Linguistic Minority Research Institute.



Teachers, Micro-Credentials, and the Performance Assessment Movement

DAN FRENCH AND BARNETT BERRY

Micro-credentials, a new form of personalized professional development for teachers, offer a unique solution to the challenge of training school staff to design and implement performance assessments.

Micro-credentials move professional development toward a more personalized learning system for teachers in which you can go at your own pace and the work is job-embedded.

– Tony Lementowicz,
Westerly (RI) High School teacher

Deeper learning outcomes for all students – and more accurate and authentic measures of them – have become the school reform coin of the realm. If this new era of performance assessment is to be successful, we need teachers to serve as assessment leaders who can help to build the literacy and capacity of every

Dan French is the executive director of the Center for Collaborative Education. Barnett Berry is the founder and CEO of the Center for Teaching Quality.

school to design, field-test, score, and refine high-quality performance tasks.

Teachers are the cornerstone of successful performance assessment initiatives. They generate, validate, administer, and score the performance assessments that are used (Tung & Stazesky 2010). Teachers need more support and training in order to fill this important role in performance assessments, yet most professional development for teachers has been found to be ineffective. Too much of the time, district central offices determine professional development focus and delivery, all but guaranteeing teacher dissatisfaction in meeting their needs and interests. A recent study points to the woeful state of our nation's \$18 billion public education professional development enterprise. The researchers found that "one-shot" workshops are the most prevalent form of professional development (Bill & Melinda Gates Foundation 2014). This one-time professional development has been found to "neither change teacher practice nor improve student learning" (Gulamhussein 2013 , p. 3). Fewer than 30 percent of teachers choose most or all of their professional learning opportunities. Only 7 percent of teachers reported that their schools have strong collaboration models (Bill & Melinda Gates Foundation 2014).

On the other hand, research suggests that the most effective professional development is contextualized to the specific needs of teachers, where they have opportunities to take ownership of their professional learning (Berry 2016). Professional development needs to be of a granular size so that teachers can engage in it during a hectic school year. Such a model often sits outside most university graduate courses, district-delivered and batch-sized professional development, and one-shot conferences.

It is within this space – placing teachers at the center of designing their own professional development, coupled with the need for teachers to build performance assessment literacy and capacity – that performance assessment *micro-credentials* come to the fore. Micro-credentials for teachers are competency-based, personalized, small-scale professional development modules that are suited for anytime/ anywhere learning and allow teachers to show what they can do, not only what they know. Micro-credentials change the face of teacher professional learning to move away from one-size-fits-all efforts to customized, just-in-time learning that leverages personal desires for professional growth.

Professional development for performance assessment literacy is uniquely suited to micro-credentialing. Both require teacher agency and collaboration, and the fact that micro-credentials can be pursued by individuals rather than schools or districts allows teachers to take the lead in scaling up to school-wide performance assessments.

SOME BACKGROUND ON MICRO-CREDENTIALS

The idea for micro-credentials began with "digital badges," which first gained recognition as a means to personalize student learning; they "are designed to make visible and validate learning in both formal and informal settings, and hold the potential to help transform where and how learning is valued" (MacArthur Foundation 2017). School districts (such as the Aurora Public Schools in Colorado) and nonprofit organizations (such as Connected Learning Alliance) are beginning to recognize digital badges, not just seat-time requirements (or a

required number of hours for courses), as markers of student achievement.¹ By enabling students to demonstrate proficiency over identified competencies (or learning targets, including dispositions such as collaboration and communication or skills as wide-ranging as set design or research skills), they are better able to track their progress in gaining tangible and usable knowledge, skills, and dispositions.

Now the personalized learning movement is reaching teachers. Over the past two years, Digital Promise, a nonprofit seeking to accelerate innovation in education, has been building an ecosystem for advancing the design, development, and implementation of micro-credentials for educators. Digital Promise has partnered with technology companies to create online professional development platforms to facilitate the process of an educator selecting a micro-credential and submitting evidence to earn it.

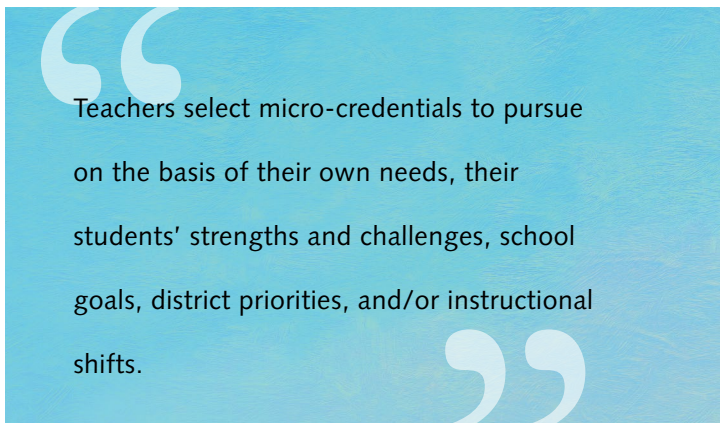
As of late fall 2016, over forty content partners have developed more than 400 micro-credentials – organized in “stacks” – to address a variety of educator skills and competencies. Micro-credentials hone in on a wide variety of competencies, from highly granular aspects of teaching (such as a unit on “wait time” issued by the Relay Graduate School of Education) to a bold brand of teacher leadership (such as the “teacher-powered” and “virtual community organizing” stacks issued by the Center for Teaching Quality [CTQ]), as well as the “performance assessment literacy” stacks issued by the Center for Collaborative Education (CCE).²

.....

1 See <http://badge.aurorak12.org/> and <http://educatorinnovator.org/webinars/why-should-you-care-about-badges/>.
2 See the online version of this article at <http://vue.annenberginstitute.org/issues/46> for links to these micro-credential stacks.

Four characteristics distinguish the micro-credentialing approach from traditional professional development systems:

- **Competency-based.** Micro-credentials focus on evidence of teachers’ attainment of actual skills and abilities, not on the amount of seat time they’ve logged in their learning.
- **Personalized.** Teachers select micro-credentials to pursue on the basis of their own needs, their students’ strengths and challenges, school goals, district priorities, and/or instructional shifts. They identify specific activities that will support them in developing each competency.
- **On demand.** Micro-credentials are responsive to teachers’ schedules. Educators can opt to explore new competencies or receive recognition for existing ones in any manner and time span they choose. They then upload evidence of proficiency using an online system.
- **Shareable.** Educators can share their micro-credentials across social media platforms, through email, and on blogs and résumés. As a result, micro-credentials can emerge as shareable currency for professional learning.



Teachers select micro-credentials to pursue on the basis of their own needs, their students’ strengths and challenges, school goals, district priorities, and/or instructional shifts.

CENTER FOR COLLABORATIVE EDUCATION'S MICRO-CREDENTIAL PERFORMANCE ASSESSMENT STACKS

1. Basic Performance Assessment

Design: design of performance assessments; design of competency-based rubrics; and performance assessment validation.

2. Advanced Performance Assessment

Design: calibrating scoring among teachers; looking at student work; assessing habits, skills, and dispositions; and using performance assessments to provide formative feedback.

3. Building and Leading a Performance Assessment Learning Community:

modeling processes for educator collaboration; facilitating performance assessment protocols in professional learning communities; and communicating and advocating for performance assessment policies at the school- and district-levels.

Each micro-credential in the Digital Promise ecosystem includes six parts: competency, key method, components, research and resources, submission criteria, and scoring rubric. Teachers assemble and upload a mini digital portfolio, which might include a video of a lesson, student work, classroom observations, teacher and student reflections, and/or other documentation of teacher learning. Trained assessors – individuals whom the issuing organization has qualified to review the evidence – examine the teacher's submission against a rubric. The issuing organization then determines whether the teacher should be awarded the micro-credential. (Since Digital Promise is still in the early stages of developing the micro-credential eco-system, the cost model for issuing micro-credentials is still under development.)

Creating performance assessment micro-credentials

In the spring of 2016, CCE and CTQ launched the Performance Assessment for Learning (PAL) initiative, with support from the Center for Innovation in Education at the University of Kentucky and Next Generation Learning Challenges. In particular, we sought to test the power of micro-credentials in promoting teacher leadership to drive adoption of schoolwide performance assessment systems that lead to personalized, proficiency-based learning and assessments for students.

We launched our initiative in partnership with the Rhode Island Department of Education (RIDE), which has a ten-year history of promoting proficiency-based education. A committee of five teachers worked with CCE and CTQ staff to design three “stacks” of performance assessment micro-credentials, with each stack containing three separate micro-credentials (see sidebar).

In the 2016-2017 school year, we brought together fifty volunteer teachers from a handful of schools, with the premise that a team of teachers pursuing performance assessment micro-credentials would be better positioned to effect school-wide change than individual teachers. These teachers came together for a half-day orientation, then worked with CCE staff individually to select their preferred micro-credentials and develop a plan of professional growth to attain them, including identifying the evidence they would collect. CTQ created a virtual community for participants to share and learn from each other.

CCE sees growing demand for its performance assessment micro-credential as states and school districts seek to build teacher capacity to transform the ways student learning is assessed. Early adopter states are

making strides toward embedding micro-credentials in their teacher certification renewal processes; for example, recently enacted legislation in Illinois allows “teachers and administrators in the state to pursue different types of professional development that can include micro-credentials” (Center for Teaching Quality and Digital Promise 2016, p. 14). Simultaneously, early adopter districts, such as Kettle Moraine School District in Wisconsin, are integrating micro-credentials into teacher salary scales and teacher leader roles. As stated on their website, “Micro-Credentials for Kettle Moraine educators . . . provide pathways to specific skills and habits that closely align to the District’s mission and goals, as well as each educator’s professional goals.”³

Early lessons: What are teachers saying about micro-credentials for performance assessment literacy?

If micro-credentials are intended to be a form of professional development that empowers teachers, then our early efforts with the PAL stacks suggest we are on the right track. At a fall 2016 forum highlighting the work of Rhode Island high schools in implementing new assessment systems, teachers piloting performance assessment micro-credentials shared their insights about engaging in learning and building a body of evidence to demonstrate proficiency over chosen micro-credentials. Several ideas emerged from listening to them:

Teachers view micro-credentials as a means to take control over their own professional development, shaping it in ways that are meaningful to them. One teacher, told us that the PAL stack helped him to “pursue his own goals,” while a second pointed out that micro-credentials “help teachers clarify what is important to them.”

Micro-credentials are viewed as a valuable means for teachers to improve their practice. A teacher noted that he had “hit a wall” with his classroom teaching. He felt like he was not getting better at his craft, and the PAL stack offered “a clear path for setting goals and improving his practice.” A high school teacher asserted, “Through engaging in these micro-credentials, I have seen the power in creating good assessments and how it improves learning for students and drives my instruction.”

Teachers value the opportunity to individualize their professional growth but also drive teaching as a collective practice. An administrator of an adult education program told us, “Micro-credentials are a perfect way to present individual learning opportunities for our professionals.” Several colleagues pointed out that their most profound utility may be in driving a collaborative process and a means to improve team-wide practice. For example, a teacher pointed out that the CCE and CTQ process of engaging colleagues in micro-credentials created “an effective formal structure for a team of teachers to ensure integrity in the process of professional growth.” Another high school teacher noted the power of the micro-credentials in defining the what and how of her professional learning community: “The micro-credential process has really focused us in a way that helps us reflect on things. For example, we are using [CCE’s] Quality Performance Assessment tools⁴ to see what we are already doing as a team and what gaps there are in our knowledge and experience.”

.....

³ See <http://www.kmsd.edu/Page/992>.

⁴ See <http://cce.org/work/instruction-assessment/quality-performance-assessment/tools-resources>.

Educators want states and districts to formally recognize micro-credentials as a credible form of professional development. While embracing the potential power of micro-credentials, educators were also keenly aware that, in order for them to be widely accepted and used, micro-credentials need to be integrated into district and state systems so that they become a viable path for teacher professional growth. As the adult education administrator asked, “Is the state going to be accepting micro-credentials as a valid credential – and can I use it for recertification?” One teacher got even more specific:

There needs to be some form of currency to incentivize teachers to use micro-credentials. This is not about seat time – it is about real learning. There are some teachers who want leadership opportunities, and micro-credentials are a way of demonstrating competencies and earning badges that schools and districts [should] value.

The teachers we interviewed are hungry for a different form of professional development – and are seeking tools and processes to spur ownership of their own learning. A recent national survey, commissioned by Digital Promise, found that nearly three in four teachers are pursuing “informal” learning (e.g., participation in online communities like the CTQ Collaboratory or Teaching Partners) that satisfies their quest to improve. At the same time, we recognize that if micro-credentials are going to gain currency as a powerful tool for teacher-driven professional development and new performance assessments, states and districts need to create incentives and opportunities to leverage the time teachers have to learn.

PAL micro-credentials and next-generation reforms

The recent passage of the Every Student Succeeds Act (ESSA) presents district and state leaders with new opportunities to rethink strategies and funding for accountability systems as well as avenues for teachers’ professional learning and growth. States have greater latitude to redefine their accountability metrics and to include a greater range of measures as part of the system of oversight and reporting. As Darling-Hammond and colleagues have noted, new accountability systems should “include annual determinations of student achievement and growth through locally designed and state-validated systems of performance assessments”

(Darling-Hammond, Wilhoit & Pittenger 2014). In such systems, a network of practitioner “assessment experts” will be needed to support schools. Each school would have two to five of these teacher assessment experts to lead faculties in the design, validation, administration, and calibration of robust, curriculum-embedded performance assessments.

With Darling-Hammond’s words in mind, to bring performance assessment micro-credentials to the fore, states and districts must take critical steps.

States need to:

- establish micro-credential attainment as a means of certification attainment and renewal;
- invest federal professional development dollars in creating well-facilitated, cross-district networks (virtual and face-to-face) for teachers to build performance assessment expertise; and
- develop incentives for districts to reallocate professional development dollars to give teachers more choice

in demonstrating their pedagogical and leadership skills via micro-credentials – with a premium on high-value competencies related to next-generation performance assessments.

Districts need to:

- create performance assessment teacher leader roles, in which teacher leaders continue to teach yet are also given time and space to build performance assessment expertise with other faculty;
- reinvent professional learning communities so that teachers have time and agency to use micro-credentials to document impact and spread best teaching practices;
- insert into salary scales the attainment of micro-credentials as a primary means of demonstrating professional growth; and
- prepare administrators to work with teachers in using the evidence from micro-credentials to spread teaching expertise.

In a relatively short period of time, micro-credentials have shown promise in enabling a more personalized, effective method of promoting teacher professional growth. Such a model is critically important in transitioning to new accountability systems that rely upon teachers on the ground to be designers, validators, and scorers of high-quality valid and reliable performance assessments. As we re-envision accountability systems to better serve student learning in complex and authentic ways, micro-credentials are an important vehicle to build the necessary teacher capacity to lead the performance assessment movement.

REFERENCES

Berry, B. 2016. *Transforming Professional Learning: Why Teachers' Learning Must Be Individualized – and How*. Carrboro, NC: Center for Teaching Quality. <http://www.teachingquality.org/transformingprofessionallearning>.

Bill & Melinda Gates Foundation. 2014. *Teachers Know Best: Teachers' Views on Professional Development*. Seattle: Bill & Melinda Gates Foundation. <https://goo.gl/vmLlrV>.

BloomBoard. 2017. "Micro-credentials for Educators," BloomBoard for Schools website, <http://schools.bloomboard.com/micro-credentials/>.

Center for Teaching Quality and Digital Promise. 2016. *Micro-credentials: Driving Teacher Learning & Leadership*. Carrboro, NC: Center for Teaching Quality and Washington, DC: Digital Promise. <https://goo.gl/aN8rSe>.

Conley, D. 2013. *Getting Ready for College, Careers, and the Common Core: What Every Educator Needs to Know*. San Francisco: Jossey-Bass.

Darling-Hammond, L., G. Wilhoit, and L. Pittenger. 2014. "Accountability for College and Career Readiness: Developing a New Paradigm," *Education Policy Analysis Archives* 22, no. 86. <https://goo.gl/tz3hOX>.

Gulamhussein, A. 2013. *Teaching the Teachers: Effective Professional Development in an Era of High Stakes Accountability*. Alexandria, VA: National School Boards Association, Center for Public Education. <https://goo.gl/2Gg6Ae>.

MacArthur Foundation. 2017. "Digital Badges," MacArthur Foundation website, <https://www.macfound.org/programs/digital-badges/>.

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

Case Study: The Capstone Project at Chelsea High School

PRITI JOHARI

Capstone projects serve as the centerpiece for an assessment system that values problem-solving and frames assessment as learning opportunities for students and teachers.

Chelsea High School educates 1,560 students in Chelsea, Massachusetts. As a gateway district, Chelsea Public Schools welcomes and educates *all* students and families. English is a second language for 87 percent of the student population; one third of our students are designated as English language learners (ELLs).

Every January and May for one week, we break from our traditional, four-periods-a-day bell schedule and the fast pace of a typical high school day. We slow everything down and ask Chelsea High School students to come in by appointment to present on a topic in which they have particular expertise: their own learning. This presentation is their capstone.

SHIFTING THE MINDSET

Chelsea High School illustrates how schools can integrate new policies and

practices into their own vision for education rather than merely follow a path of compliance with district and state mandates. In 2012, we decided to develop our own authentic formative assessment system aligned to the Massachusetts Frameworks and state standardized tests (MCAS) in 2012. We strive to integrate performance assessment into everyday work, including teaching, evaluation, coaching, training, and faculty meetings. We have shifted our mindset to viewing assessment as a learning opportunity in and of itself – for both students and teachers – and have adopted the phrase, “learning by doing.” For example, the science lead commented:

In the past, I used to think that students needed to completely understand certain scientific facts first, in order to be able to solve authentic scientific problems. But in practice, I seldom let students try those scientific problems, because

Priti Johari is the principal of Chelsea High School in Chelsea, Massachusetts.

students never showed a complete understanding of the relevant facts. Once we switched to performance assessments, I saw that students are capable of learning scientific facts at the same time as they are solving a problem. As a matter of fact, I think that their learning is enhanced: if students see that they need to solve an engaging, authentic problem, they will be more motivated to learn the facts needed to solve the problem.

This assessment system has been a long time in the making. Our transformation is a result of sustained work around a coherent vision as well as a deliberate investment in our capacity to purposefully collaborate across lines of authority. The process involved many steps: building structures and skills for collaboration, articulating a “Vision of the Graduate,” unpacking the Massachusetts Frameworks, writing lesson plans and designing performance tasks, sharing those plans, giving each other feedback, calibrating on common rubrics, and collectively examining student work. Every step has involved individual teachers, department lead teachers, instructional coaches, administrators and curriculum coordinators in development of the overall assessment system as well as key curricula such as department rubrics and performance assessments.

ASSESSMENTS AS LEARNING OPPORTUNITIES FOR STUDENTS

Throughout the year, students work on using quarterly performance assessments (QPAs) to demonstrate their understanding in courses ranging across English Language Arts, Pre-Calculus, Painting and Drawing for Realism, and World History. Capstone presents students with an opportunity to review and revise their QPAs, cement their learning, *and* reflect on

who they are as learners. Using their work, revisions, and rubrics as evidence, students articulate and analyze their growth throughout the school year, noting specific challenges and triumphs.

During one presentation, a sophomore enthusiastically explained why he performed poorly on a first-quarter history assessment, and how he used his rubric and teacher feedback to improve his writing and analysis.

Another student detailed his process for decoding a new piece of music and how he might, going forward, apply the same technique to his algebra course. Performance assessments give students the opportunity to apply new content knowledge and skills. Capstone allows space for “meta-cognitive learning,” where students can step back, reflect, and modify their approach.¹

Performance assessments are different from traditional tests. A CHS alum shared:

A test shows if you know obvious facts, but when I did something like [the historical argument essay], it actually made me take the time to concentrate. I was able to really think and really understand and dig through all these [historical] documents so I could really understand.

The inclusion of reflection and revision as part of the cycle adds rigor, allowing for a second, meta-cognitive level of learning. A current senior commented:

QPAs are good for me as a student, because I am not only developing skills in class but also developing skills as a person. Capstone gives me the opportunity to reflect on myself as a person.

.....
¹ For more on meta-cognitive learning, see <https://lincs.ed.gov/programs/teal/guide/metacognitive>.

TEACHER LEARNING AND COLLABORATION

QPAs and the capstone project are also learning opportunities for CHS teachers. Successful implementation of QPAs or capstone does not come from purchasing materials, programs, and new assessments; it comes through sustained conversation with and among teachers grounded in student work. A learning system where assessment and instruction are integrated requires an investment in job-embedded professional development to support the teacher learning to effectively implement the new standards.

We are creating a collegial space and building a school culture through content-based Professional Learning Communities (PLCs), which meet weekly during the school day and are led by instructional coaches and administrators. Within the PLCs it is not only okay for teachers to put their work on the table and engage in public learning; it is required. Teachers engage in co-construction of lessons or QPAs, lesson rehearsals, and data conversations rooted in student work to explore what proficiency in the standards means, what students understand and can do, and what the task design allowed students to demonstrate.

For example, in a recent PLC, a group of algebra teachers grappled with how to design an exit ticket that would demonstrate if students understood that all points that satisfy a particular equation are on the graph of the line. They wanted to design a question that had multiple solution pathways and checked for conceptual understanding, not just mathematical fluency. PLCs allow teachers, instructional coaches, and administrators to foster a culture of collaboration, inquiry, and public learning across grade-levels, content areas, and traditional authority lines.

The process of creating departmental rubrics has been a critical step across all content areas. The visual arts lead recounted:

As a group, we were able to identify the most valuable artistic habits we wanted our students to have. We realized that skills in reflection, criticism, and process were often previously marginalized at the expense of artistic technique. Now, because the rubric requires us to evaluate the other types of skills, they have been integrated much more fully into instruction and daily practice.

The teacher added that the benefits are significant for teachers and students:

Students who are able to more independently follow an artistic process and think critically make better artwork. Plus, these skills are transferable to other subjects.

When considering common rubrics and common assessments, some teachers may worry that this kind of collaboration may have a negative impact on their autonomy or creativity in the classroom. However, having shared endpoints has not resulted in uniform units or lesson design; every teacher must be responsive to the students in front of him or her to determine how to help them reach these common goals. The work is often perceived as more collective because there are common expectations and language and, in turn, shared dilemmas and solutions developed and owned by the entire team.

In Chelsea, the work has been empowering for teachers and students. As one teacher noted:

I also appreciate how the specific language in rubrics helps students with goal setting and provides concrete feedback. They are able to clearly see why they are being scored

at a “developing” level rather than “proficient” because the rubric language is so clear and specific.

As a result, students can talk about themselves and their academic knowledge and soft skills in a way that puts them in control of their learning and futures – and allows them to see a way forward. One CHS senior said:

Performance assessments are different than regular tests because actual tests do not ask how you prepared for it. [A] performance assessment gets in depth, it talks about how you prepare for the test and what skills you used. Performance assessments impacted me as a student by getting to know what my flaws were.

The work of creating and completing meaningful performance assessments is not easy and requires shared commitments from both teachers and students. Even with that in mind, the ELA lead reports:

I’ve seldom had a student who didn’t seem to genuinely enjoy talking about their learning, particularly the accomplishments they’re most proud of. I believe that capstone is a renewing experience for teachers and students alike; it allows us to think about what we’ve accomplished (and how) and to set goals for future teaching and learning. Personally, I’ve found that it helps me to get to know students on a deeper level as learners and as people, and it has helped to build a sense of community and support in my classes.

The shift in our understanding of students and actual student learning also, necessarily, leads to shifts in teacher learning. As one teacher reflects:

Recently I looked at some of the reflection activities and evaluation methods I used in my own instruction in the years before redesign. It was clear to me then how much better my teaching has become. I am asking more difficult questions, holding students more accountable for high-quality responses, and giving them more robust and better-quality feedback.

For more about Chelsea High School, visit <https://www.chelseaschools.com/cps/schools/high-school.htm>.

“I believe that capstone is a renewing experience for teachers and students alike. Personally, I’ve found that it helps me to get to know students on a deeper level as learners and as people, and it has helped to build a sense of community and support in my classes.”

— Chelsea High School teacher



District, Union, and Community Collaboration: Massachusetts Consortium for Innovative Education Assessment

DIANNE KELLY AND ERIK FEARING

Consortium leaders discuss how their model – based on collaboration among districts, teachers unions, and community organizations – aims to change the way school quality is assessed.

Since 1996, Massachusetts’s accountability system has been defined by a single standardized test. While Massachusetts has received accolades for its high scores on the National Assessment for Educational Progress (NAEP) and Program for International Student Assessment (PISA) tests, absolute results from

standardized tests tend to correlate strongly with family income and parental education (Reardon 2011). The state’s aggregate scores mask significant inequities; Massachusetts ranks in the bottom third of states with the largest achievement gaps by race, income, and language. Additionally, the narrow focus on one set of metrics

Dianne Kelly is superintendent of Revere Public Schools. Erik Fearing is president of the Revere Teachers Association and co-chair of the Massachusetts Consortium for Innovative Education Assessment.

– test scores in three subject areas – has incentivized narrowing the curriculum and teaching to the test, particularly in urban districts that serve diverse students, due to the pressure to avoid being designated as an underperforming school or district.

The Massachusetts Consortium for Innovative Education Assessment (MCIEA) was founded in March 2016 by a group of school districts, teacher unions, partner organizations, and a key state senator, with the goal of creating a new multiple-measures accountability system. This system was founded on the belief that there are richer assessment methods than a single standardized test to truly assess student learning and school quality, methods that can provide a deeper, more dynamic understanding of students and schools.

The consortium is creating a multiple-measures school quality dashboard in the areas of teachers and the teaching environment, school culture, resources, student learning, and civic engagement and well-being. The primary means of assessing student learning will be robust teacher-generated, curriculum-embedded performance assessments. The completed dashboard will provide parents, students, educators, community, and policymakers with a comprehensive portrayal of how a school is progressing across all the areas that contribute to students' social-emotional and academic growth. Rather than one single score, the dashboard will show the areas in which a school is doing well and those where more progress is needed, providing more complete and accurate data to use in determining improvements that need to be made. Such a dashboard eliminates the need for single scores, ratings, and levels that currently exist merely to sort students and schools.

The consortium's governing board consists of superintendents and teacher

union presidents of consortium districts, with the Center for Collaborative Education (CCE) providing coordination support and facilitating the performance assessment initiative, and a team from the College of the Holy Cross facilitating design of the school quality dashboard.

CCE's executive director Dan French sat down with Dianne Kelly, superintendent of the Revere Public Schools (a consortium district), and Erik Fearing, Revere Teachers Association president and MCIEA co-chair, to discuss the consortium.

What is the primary goal of the Mass Consortium?

Erik: The biggest impact would be a change in the state education culture from a focus on punitive accountability and multiple-choice testing to a holistic recognition of student knowledge and the value of schools and districts within their communities. I want to see teachers' professionalism recognized and have them be involved in the formal assessment of students and schools. We need to move away from the extreme focus on one limited style of assessment – standardized testing – and on limited subjects, and recognize the broad value of schools across all academic subjects as well as non-academic areas.

Dianne: Right now, in this state and across the country, there is a false narrative about the efficacy of public education. We constantly hear in the media about failing schools and schools that aren't meeting the accountability standards in Massachusetts. There are many powerful voices in the commonwealth that support charter schools under the guise of our "failing public schools" and the idea that the public schools don't innovate or meet individual student needs. Those are untruths about what is really happening in public education today.

The Revere Public Schools is a perfect example of an urban school district in which we have high levels of poverty and a diverse student population, and yet our schools are succeeding. But there is no one beating down our doors to do a story about that; instead, they are talking about how bad the public schools are. So we need to find alternative ways of demonstrating that our public schools are, in fact, succeeding.

The idea isn't that we abandon accountability. Rather, we want to abandon having one test be used to make judgments about a student or school. A number on a standardized test does not speak to a student's unique needs; single scores do not adequately describe the kinds of successes that a particular child may have had in school. For example, for a student with an interrupted formal education and the social-emotional effects of living in a war-torn country, success might be to spend an entire day in school with classmates and not have an emotional meltdown. But we don't get to talk about that when all we talk about is a student's score on PARCC or MCAS (Massachusetts's standardized test).¹ MCIEA creates an opportunity to assess schools holistically for how they are able to help students achieve instead of looking at narrow, non-descript, decontextualized, single test scores.

Why did you want the Revere Public Schools to join MCIEA?

Dianne: Our current accountability system highlights a narrow area of focus. We should value academic disciplines in addition to mathematics, English language arts, and science. We seek to create a different assessment system that incorporates performance assessments so that, for example, students can articulate their knowledge of mathematics through an artscape they create. We want to enable stu-

dents to show us what they know and can do in multiple ways instead of pigeonholing them into "show me this one way to do this, which is the way I want it done, and if you can't, you're a failure." MCIEA creates an opportunity for students to express themselves through learning that engages them.

Erik: One reason I am interested in MCIEA is to get a better measure for districts. The current accountability system² is punitive for districts with high percentages of low-income students, almost guaranteeing that these districts will be at the bottom of the list. We are looking to create measures that everyone believes in other than scores from multiple-choice exams.

How does a district become a member of the consortium?

Dianne: From the very beginning, we wanted to be sure each district was represented by a superintendent and local teachers union president. The truth is, regardless of what the superintendent wants to implement, whether or not it happens at the classroom level is up to the teacher. So if the teachers are not on board from the very beginning in making decisions and shaping the program, then it would be less likely to succeed.

Erik: This has to be a grassroots effort, something that teachers believe in.

-
- 1 These standardized tests are given to Massachusetts students from third through twelfth grade. PARCC (Partnership for Assessment of Readiness for College and Careers) tests cover English language arts and math; MCAS (Massachusetts Comprehensive Assessment System) tests are given in English language arts, math, and science.
 - 2 Massachusetts' public schools are sorted into levels from high-performing to lowest-performing based predominantly on student scores on the state's standardized test; test participation rates and graduation rates (at the high school level) are minor factors in determining levels.

Having union leadership at the governing-board level gives credibility to empower teachers in the decision-making process. It's not easy to get a relationship of trust between a superintendent and teachers union president where you can have genuine collaboration. There is a power disparity between them. So having both parties on board and having the broader membership across the district behind those people is crucial to staying in.

What is the balance of the work between school quality and performance assessment?

Erik: In the beginning, each district entered the consortium for different reasons, and some weren't sure they were going to commit to both paths. In the end, though, every district committed to engage in both parts.

Dianne: Both avenues of work are extremely important. In Revere, it made sense from the very beginning to get in on both of them and not feel left behind on one or the other.

What is the school quality measures work going to look like?

Erik: We want to completely overhaul the measures that are used in determining school quality. We are asking stakeholders within each district – including parents, teachers, students, and administrators – what it is they think makes a great school. We will use the answers to build a school quality measures framework, and then gather available administrative data and develop surveys for different stakeholders, such as gauging how newcomer English language learners – students and their parents – feel welcomed in school.

Dianne: Revere High School has recently received two national awards because of our work in welcoming a diverse student population to the

building and meeting the needs of our diverse students. The process leading to the awards involved site visits with teams from multiple states who dissected the curriculum and enrollment, examined whether students of color were well represented in Advanced Placement classes, observed the quality of the advisory period, analyzed discipline data, and gauged the relationships among students, between adults and students, and among adults. They dissected the entire school. In both cases, Revere High School was the only gold school winner from New England. Hugely impressive.

Yet, in the Massachusetts accountability system, Revere High School is at the 22nd percentile in performance, which is based almost completely on a single standardized test average. With a handful of lower MCAS scores, Revere High School would have been classified an underperforming high school and in need of intervention by the state. Clearly, the time is right for us to be talking about why Revere High School is determined to be outstanding on the national level but is deemed as borderline in trouble at the state level.

What is the thinking behind moving from an external testing company to teacher-generated performance assessments in order to assess student learning?

Erik: Standardized tests are efficient at providing scores and a ranking, but they are not very effective at actually assessing learning, knowledge, and skills. With performance assessments, teachers are developing, administering, and scoring tasks. Teachers, working together, will be the ones who examine student work and determine whether it meets the proficiency benchmark. This process recognizes the professionalism of teachers. Instead of devoting so much time away from the curriculum to taking these external state tests,

we'd much rather have these well-vetted, thought-out performance assessments where we are getting better information about student knowledge and capacity from an assessment that is part of the curriculum. Students will learn something from taking the test, and we won't lose instructional time.

Dianne: What is exciting for me is to get at the question, "What is the purpose of assessment?" If our purpose is to assign a numerical value that determines the rank of a school against all other schools that serve students at a particular grade span, I'm not sure who that helps. The Massachusetts

are going to articulate their knowledge and skills advances their learning. Engaging in that type of performance assessment is a much more valuable use of time, effort, energy, and resources than to associate a particular number with a particular student and with a particular school.

What will implementation of the performance assessment work look like in MCIEA districts?

Erik: This first year, we've started with professional development for creating and piloting performance assessments. There are thirteen schools in this first cohort. Each school has determined the grade level and subject areas represented on the lead teams. Over four years, we expect to engage every school in each consortium district in the performance assessment work.

Dianne: I envision a time down the road when using performance assessments in class is just a routine part of what we do. And teachers meeting in teams to review student work and refine performance assessments is a cyclical thing we do in order to determine who is achieving understanding of the curriculum, who needs additional assistance, and what additional assistance we need to give them.

Will participating districts be exempt from MCAS and from the underperforming designation?

Dianne: Not while we are building our accountability system. Ultimately, though, the goal is that we are able to provide robust data about student achievement and school progress; at that time, we will sit down with the state and request them to apply for a federal waiver that would exempt participating districts from MCAS.

"Assessment should be a way to inform teacher practice and help students understand their progress, and giving students choice about how they are going to articulate their knowledge and skills advances their learning."

— Superintendent Dianne Kelly,
Revere Public Schools

accountability system automatically says that 20 percent of schools at each grade span have to be failing. It doesn't matter how good those schools are; somebody has to be in the lowest 20th percentile and labeled as failing. I don't know how that system speaks to what our kids know and are able to do. It's nonsensical, really.

Assessment should be a way to inform teacher practice and help students (and parents) understand their progress. Giving students choice about how they

Erik: We would like to see the current performance rating, single number scores that are given to schools and districts, go away. There is a lot more to a school than is measured on MCAS.

How will participating districts ensure technical quality?

Dianne: The fact that teachers are meeting in cross-district groups to vet the assessments and score student work will contribute to ensuring technical quality. Teachers are going to receive substantial professional development on how to write an effective performance assessment with rubrics and how to assess appropriately. What is important for us to work on as a district is how teachers can work together to make sure that implementation throughout their schools is of high quality.

Erik: Once we have draft tasks from multiple districts, cross-district teams of teachers will be able to look at them and get a second set of eyes on them. We will also be partnering with the Center for Assessment to ensure the right technical quality measures are in place.

How does Revere envision providing adequate professional development time to implement performance assessments?

Dianne: In order to be selected for this work, each principal and school leadership team had to agree to devote a good chunk of professional development time to building school-wide faculty capacity to create, validate, and score performance assessments. That will help pollinate the work across a school. As well, in Revere we allow teachers to select and sign up for ten hours of professional development in any area that interests them; working on performance assessments will be

one area in which they can choose to focus.

Erik: The school leadership teams participating in the performance assessment institute have spent time putting together implementation plans on how to build capacity within their own buildings. And we'll have another cohort of schools going through that same process next year. In the long term, there is a question of ensuring we create a high level of expertise. We have a quality amount of school-based collaborative time, so teachers can work on and share practice around performance assessments in professional learning groups.

Why do you think the state legislature supported the consortium's work by including a budget line item to support MCIEA?

Dianne: I think even our legislature understands that it is time, almost twenty-five years after MCAS was introduced, to reflect on what we have learned and set new, loftier goals for our schools and our students' achievement. Over recent years, we have been able to identify effective best practices in instruction and assessment. We're well positioned to move forward into this new era of assessment and look at the purpose of assessment differently than we did twenty-five years ago.

Erik: The 1993 Education Reform Act set a vision for a multiple-measures state assessment. Unfortunately, the test that was created did not hit the mark; we ended up with largely a multiple-choice test. We are looking to capture that spirit again, to create an accountability system that measures everything we want to measure and that the legislature wanted to measure in 1993. We have a lot more capacity now to work toward that goal.

What challenges have you faced in launching the consortium?

Erik: It has been important to get buy-in from all union members. The first step has been getting teachers enough information to ease the anxiety of an unknown initiative. We are talking about big changes in assessment practices, and change is hard. Many teachers feel that MCAS was misused and worry that MCIEA assessments will be similarly misused, so a big piece is communicating to all teachers so they have a full understanding to be on board. Being responsive to people's concerns meant slowing down our start-up a little bit, but doing so has positioned us well for moving forward.

Dianne: I think there is a historical context where teachers sometimes think that new initiatives are coming down from on high and the union doesn't have a say in what it's going to be or even whether they want to do it. We had to make it clear that joining MCIEA was a joint district-union initiative and that teachers would have a say. Whenever we talked about MCIEA to teachers, Erik and I talked about it together. That made a difference.

To sum up, what's the message you most want to convey about the work of the consortium to the public and policymakers?

Dianne: Believe in us. We need less testing and more assessment *for* learning rather than assessment *of* learning.

Erik: Schools aren't failing – that's just a narrative that policymakers decided to write and have stuck to for a long time to maintain a certain power structure.

For more on the Massachusetts Consortium for Innovative Education Assessment, see <http://mciea.org>.

REFERENCE

Reardon, S. F. 2011. "The Widening Academic Achievement Gap between the Rich and the Poor: New Evidence and Possible Explanations." In *Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children*, edited by R. Murnane and G. Duncan. New York: Russell Sage Foundation Press.



Identity Affirmed, Agency Engaged: Culturally Responsive Performance-Based Assessment

RICARDO ROSA

Performance assessments must be culturally responsive in order to truly serve the needs of students from all backgrounds.

To be hopeful in bad times is not just foolishly romantic . . . If we see only the worst, it destroys our capacity to do something. (Zinn 2004)

I am not comfortable with categories of identity because I have witnessed the way that they are utilized to arrest the mind, detain the spirit, and even liquidate a people. Moving through the discomfort, I have come to embrace an identity as an “activist-scholar” or “scholar-activist” and to realize that identity is always a process of negotiation between how we see ourselves and how we are seen.

Ricardo Rosa is an associate professor in the Department of Educational Leadership at the University of Massachusetts, Dartmouth.

For the past few years, I have been active in the resistance against high-stakes standardized testing. I have always been vocal and have signed letters against the misuse and abuse of standardized testing. This defiance was triggered by the pressures placed on my son to take the Massachusetts Comprehensive Assessment System (MCAS) standardized test, despite the onset of the flu. As I said then, and as I continue to believe, my children are not my only concern. The regime of high-stakes testing is deeply disturbing

Institutions, including schools, are designed in the image and interests of those who rule, so we must be very cautious about recreating an educational reform environment where people of color and the poor will continue to be marginalized.

and inhumane for all. Educators standing before students and communities touting the virtues of these tests should be ashamed. And if they are aware that they are wrong but remain silent, they are complicit in educational malpractice.

Since then, I have organized community forums to translate educational research to the lay public and to advocate for those who choose to opt out of high-stakes testing. I have been active in local, regional, and national organizations contesting not only the test but, more broadly, the corporate agenda in education. I seek deeper thought about what is possible, to

simultaneously struggle against oppressive measures while elevating liberatory practices. To paraphrase Eleanor Roosevelt, I hope to join forces with those who are lighting candles and not just cursing the darkness.

A great deal has already been stated about performance-based assessments in this issue. It has been defined in various ways and examples have been shared. I wish to engage the topic from a different angle. W. E. B. Du Bois once asserted, in light of a great deal of dialogue regarding systemic attacks on people of color, that “a system cannot fail those it was never meant to protect.” If we begin, as I do, from the perspective that institutions, including schools, are designed in the image and interests of those who rule, we must be very cautious about re-creating an educational reform environment where people of color and the poor will continue to be marginalized. If performance-based assessment is considered in the same frame as current testing regimes, which is entirely possible, it becomes just another reform fad (I don’t mean to suggest that performance-based assessment is a new one) that re-inscribes the power of systems of categorization and the conferring of rewards to those who are already materially, racially, and culturally privileged. From this perspective, performance-based assessments become another repressive surveillance technique in the lives of children and adolescents. The point is to expand performance-based assessment by rethinking its boundaries. We need to appropriate it and simultaneously remake it so that it’s not a space of colonization, a practice of arresting minds through curricular and social control, but a space that allows us to speak and act beyond the boundaries of domination.

These assessments remind me of the story of Amira,¹ a six-year-old first-grader who immigrated to the United States in September 2016 and now attends a public school in a Massachusetts school district. I came to know her story through conversations with a refugee resettlement social worker, a school official involved in her case, and my own research. She was born in Syria under conditions that can only be described as an epic tragedy. Her family fled to a refugee camp in Jordan when she was three, and they were on a waiting list for asylum for three years. The refugee camp from which Amira hails is widely considered dangerous, with conditions that are especially precarious for women and children, including frequent altercations, sexual assault, and a general lack of safety. Medical services were provided only at the very basic level, and low-quality food triggered additional health care concerns while Amira was in the camp. At an early age, she witnessed a large number of civilian casualties and experienced the constant state of panic related to indiscriminate bombings. Her parents also witnessed these horrific scenes and suffered greatly as a result of U.S.-led sanctions against the Assad government. Amira did not have access to a normative school while in the refugee camp.

Teachers at Amira's new school have noticed that when she transitions from the classroom to any area of the building, she leans and brushes against walls. She's likely exploring her environment and sensing the difference between leaning and brushing against a tent (a common past behavior in the only dwelling she has known) and a more durable surface. She has difficulty walking in straight lines. She constantly wanders about the classroom and has been described by her teacher as

.....

1 Amira is a pseudonym.

impulsive. If she sees something in the room that she wants, she immediately attempts to retrieve it. Several teachers have claimed that she likely has a neurological disorder and perhaps a dis/ability related to motor development, because of her inability to hold scissors and appropriately cut paper. Part of the problem, it seems, is that her teacher is frustrated about not being able to communicate fluently with her, given that her native language is Arabic.

Teachers around Amira did not take the time to inquire about her history. Had they explored her background, they would have understood that the political, economic, and social conditions that she was exposed to would be disabling for anyone.

Her rate of English acquisition from September to November 2016, despite never having access to formal schooling, was spectacular. She could articulate primary colors, numbers, clothing items, and simple sentences in English. The greatest evidence of her intellectual acquisition and cultural immersion came in October, when she learned words such as pumpkin, Jack-o'-lantern, and Halloween. After hearing that several teachers had concerns about her progress, her parents were quite surprised. They felt that she was making outstanding progress with her acquisition of English. In addition, she has gradually learned to refrain from leaning and brushing against walls, yet it seemed that the school's narrative of this behavior did not change.

Amira was, in fact, engaged in a natural process of cross-cultural scaffolding that invariably incorporated theorizing her relationship to her new environment. She observed a classroom alive with color and manipulatives. The disposition of the room invited exploration, and she seized it. These are luxuries that were

non-existent in the refugee camp. The problem was that her teacher and the organizational behavior of the school was fraught with structures, routines, and schedules. Her exploration had to be pursued within the context of that structure and culture of efficiency. And, of course, the culture of efficiency militated against the development of a strong presence of mind with regards to cultural difference. By “difference,” I am not suggesting that the analytical focus be placed solely on Amira’s life history, but also on the ways in which our own cultural conditions get so normalized that the assumption becomes that this is the correct standard by which all others should be measured. Amira was, in short, not atypical given her history and the contexts she was navigating.

A more flexible, culturally responsive system of assessment could have captured Amira’s progress and encouraged her to continue to heal and learn. If the organizing principle behind performance-based assessments solely concerns an evaluation of the student in relation to a set of predetermined standards, many children like Amira will be at a disadvantage. Students who are indigenous, African-American, immigrant, LGBTQI, and dis/abled (to name a few markers) will be subjected to a process that requires them to adopt values and dispositions that negate their own identities. Amira’s story is a testament to the fact that children and adolescents are perpetually engaged in theory outside of the boundaries of what teachers might deem performance-based assessment. She was engrossed in imaginative play and creating her own experiential curriculum at the boundaries of cross-cultural contact.

It is therefore imperative that we continually assess and reflect on how our objectives, outcomes, and forms of evaluation relate to or negate the history of the child and the cultural,

social, political, and economic context from which the child is coming. It is also important to understand that this assessment should be perpetual. Amira, for example, is dealing with another set of challenges related to her refugee status in a political environment that demands her erasure. She may not have a language to name the condition and experiences, but then again, many adults are unable to articulate it as well. What responsibility do we bear to identify and address the multiple challenges that students like Amira face, through structural and political barriers that our systems set up?

Performance or portfolio-based assessments, seamlessly integrated into curriculum and instruction and offering learners and educators plenty of opportunities to self-reflect, are decidedly powerful. The example of the New York Performance Standards Consortium is perhaps the best illustration (see article by Robinson and Cook in this issue). Yet learning outcomes are not the only positive aspect of these assessments. Performance-based assessments, at their best, assist us in the reconnection with youth and their full being. Scholar-activist Vajra Watson has been able to pair educators with community-based spoken-word artists through processes that allowed educators to develop a greater presence of mind about the material conditions of students and their cultural contexts (Watson 2013). This example of a community-based professional development of teachers has expanded to include multiple classrooms in multiple schools in Sacramento, California.

Performance-based assessments are not only necessary for engaged teaching and learning; they are imperative for life in any society committed to the ongoing democratization of civil society. They are essentially about building the dispositions and human connections essential to deep democracy.

Schools tend to be highly undemocratic spheres where various oppressive ideologies converge. A democratic political system cannot come to fruition if the institutions of that society are undemocratic, anti-democratic, or fail to (re)create the structures and conditions that lead to further democratization. Democracy flourishes when democratic cultures are the norm. Performance-based assessment, pursued correctly, is not just a technique or routine, but essentially a way of being that allows democracy to be lived on the bones.

To be more vigorous, however, performance assessments must be critical in a dual manner: in the sense of provoking imagination and in unmasking and intervening in relations of power. In the case of Amira, deeper reflection on the part of teachers concerning the relationship between knowledge and power could have unfolded. Why were teachers seeing some of her actions as deficits instead of consciously searching for the ways that she negotiated the adjustment to a new social situation? Meaningful exchanges with Amira, perhaps recorded to allow teachers to iteratively analyze her meaning-making process, might have led to an understanding that strengths were being exhibited and that those strengths should be integrated into formal assessment.

Greater thought could also be given to Amira's current experiences, and formal lessons and assessments might be designed on virtually any topic in order to expose academic content while allowing her to further explore her conditions. For example, her story of geographic movement and space is critical to her; how might her teacher utilize this knowledge to help her arrive at a greater understanding of her story? Knowing, accepting, and understanding our stories are fundamental acts in acquiring power. A simple introduction to a world map that allows her to understand her

family's movement is a first step. And this kind of mapping activity can involve a wide range of competencies: math, language arts, art, science, and geography.

If performance-based assessment is going to be of deep value, it must integrate not only the ordinary stuff of curriculum, but also the extraordinary as exemplified by life histories like Amira's – and the extraordinary is all around us. Teachers must learn to see what's there, what's not there, and theorize what should be there. Whose knowledge and experience are licensed in the very formation, implementation, and development of the assessment? What are the rules of power circulated organizationally that privilege some and marginalize others? What are the relationships between the material conditions in which youth live and performance?

How might a deeper understanding of these asymmetries provoke a more thoughtful approach to performance-based assessment? How do ideas about students' cognitive or motivational "deficiencies" and family or cultural deficits factor into the ways in which we (re)construct these assessments? We must be provoked to a presence of mind or an ethnographic eye/I (Ellis 2003) where every uncovering of what students truly know is also an unraveling of the boundaries of our own identities, knowledge, and comfort. If what we are looking for in performance based assessment is validation of our own ways of seeing and being, what we are in fact reproducing is cultural oppression.

Performance-based assessment is unquestionably superior to the instrumental rationality of high-stakes standardized testing and the audit culture that testing regimes inspire. It is more likely to engender opportunities to witness the un-measurable: vision, imagination, and compassion. But it

must also invite students like Amira into a culture of questioning in which her identity is fully embraced and where she is able to find a way to channel her learning and emotions into positive projects that allow her to be a subject of history and not an object to be worked on. It must minimize the distance between learning and everyday life – a gap so often dismissed in schooling. It must be critical in stimulating engaged learning, and it must be critical in provoking social agency.

Plenty of examples of performance-based assessments exist. The key, I suspect, is not to pursue it as a method, but as a process that is always defined contextually. Each school, community, and socio-historical context provides unique opportunities to re-imagine these assessments and curriculum/instruction more broadly. For example, I have spent significant time at a school in the South Bronx called The Cornerstone Academy for Social Action.² The school has been incredibly effective at examining poverty across the curriculum. In my last visit, I witnessed a breathtaking discussion between seventh graders on James McBride’s *The Color of Water*. It reminded me of my own time as a student in doctoral seminars; the discussion was just as intense.

At the end of the series of lessons across subject areas, students spent time creating hip-hop videos and art as a culminating assessment. The Bronx, of course, was the birthplace of hip-hop, so the connection to context was powerful! Furthermore, the staff and students have organized marches to protest community-based traumas such as the news of a grand jury’s decision not to indict a police officer in the fatal shooting of Michael Brown.

.....

2. See <http://www.casamiddleschool.org/>.

The school has also organized protests against budget cuts. Students were not being prepared for life in a democracy in some distant future; they were living democracy in the moment. And, contrary to popular belief, standardized test scores did not suffer. In fact, the school has one of the highest performance rates on standardized tests in the district.

Even within the context of testing regimes, the opportunities for a more hopeful education are abundant. But if performance-based assessment is going to be of any value, it must be situated within a comprehensive process that animates educators to move beyond our own comfort zones and assist us in being more self-reflective, equity minded, and socially engaged.

REFERENCES

Ellis, C. 2003. *The Ethnographic I: A Methodological Novel about Autoethnography*. Lanham, MD: Alta Mira Press.

Watson, V. 2013. “Censoring Freedom: Community-Based Professional Development and the Politics of Profanity,” *Equity & Excellence in Education* 46, no. 3:387–410.

Zinn, H. 2004. “The Optimism of Uncertainty,” *The Nation* (September 2). <https://www.thenation.com/article/optimism-uncertainty/>.



Annenberg Institute for School Reform

AT BROWN UNIVERSITY

Annenberg Institute for School Reform
Brown University, Box 1985
Providence, Rhode Island 02912

Web: www.annenberginstitute.org

Twitter: [@AnnenbergInst](https://twitter.com/AnnenbergInst)

Facebook: www.facebook.com/AnnenbergInstituteForSchoolReform

Non Profit Org.
U.S. Postage
PAID
Providence, RI
Permit #202



This book is printed on Environment® Paper. This 100 percent recycled paper reduces solid waste disposal and lessens landfill dependency. With this project, the following resources will be saved:

- **2,341 lbs of wood**, which is equivalent to **8 trees** that supply enough oxygen for **4 people** annually.
- **3,419 gallons of water**, which is enough for **198 eight-minute showers**.
- **2mln BTUs of energy**, which is enough energy to power the average household for **10 days**.
- **208 lbs of solid waste**, which would fill **46 garbage cans**.
- **710 lbs of emissions**, which is the amount of carbon consumed by **8 tree seedlings** grown for **10 years**.

