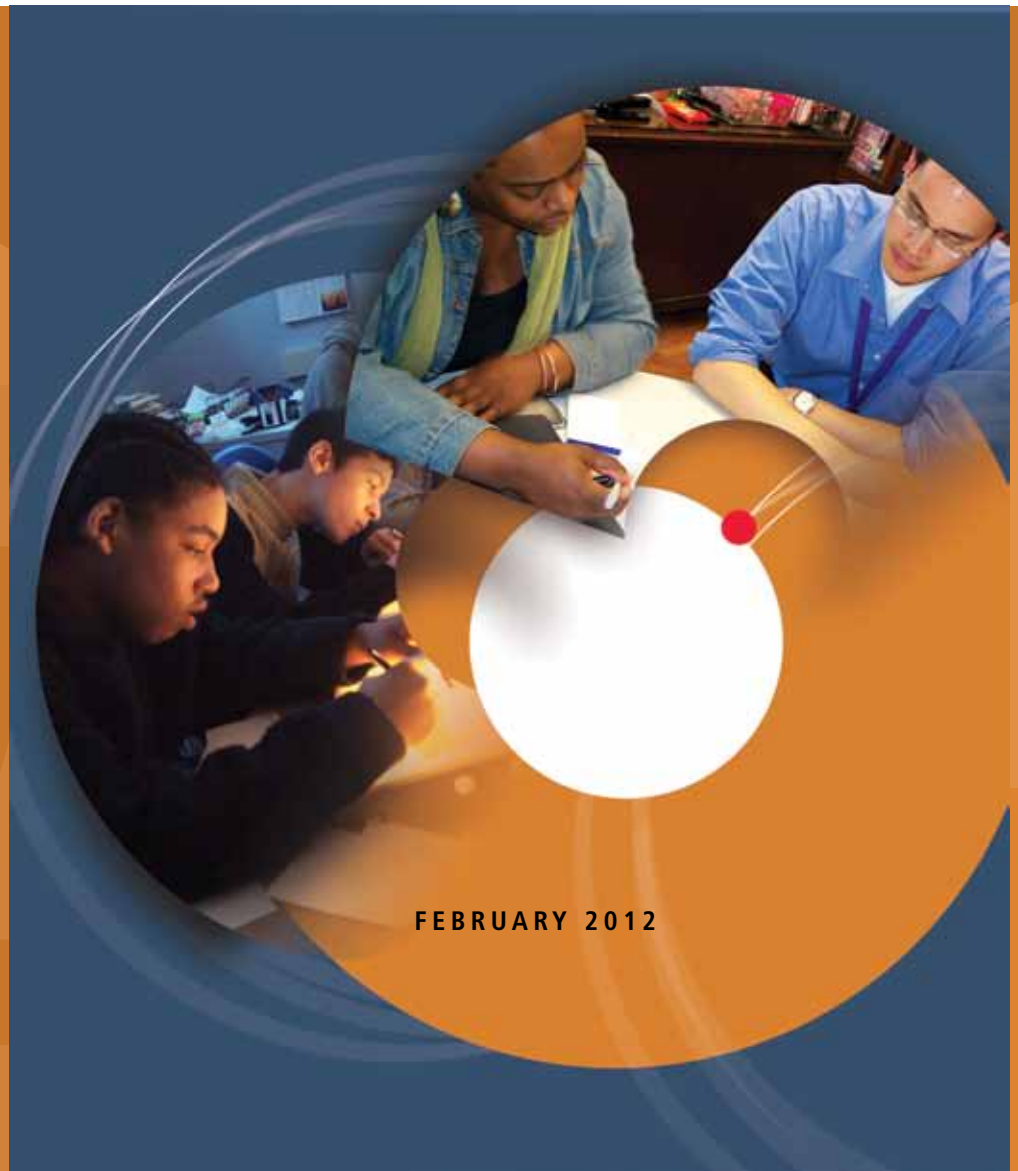


QUALITY PERFORMANCE ASSESSMENT: HARNESSING THE POWER OF TEACHER AND STUDENT LEARNING

Christina Brown and Pascale Mevs



FEBRUARY 2012



TABLE OF CONTENTS

| | |
|---|----|
| INTRODUCTION | 1 |
| A FRAMEWORK FOR HIGH QUALITY PERFORMANCE ASSESSMENT SYSTEMS | 3 |
| Fundamental Elements of Quality Performance Assessment | 3 |
| Essentials for Success | 4 |
| THE QUALITY PERFORMANCE ASSESSMENT FRAMEWORK IN ACTION—THREE STORIES FROM THE FIELD | 6 |
| I. Quality Data Analysis: Cape Cod Lighthouse Charter School Assessment Validation Sessions | 6 |
| II. Quality Task Design: Fenway High School Junior Review, Senior Institute and Senior Position Paper | 12 |
| III. Quality Aligned Instruction: Pentucket Regional School District Habits of Learning | 18 |
| CONCLUSIONS: KEY THEMES AND LESSONS FROM STORIES FROM THE FIELD | 23 |
| STEPS TO QUALITY PERFORMANCE ASSESSMENT | 26 |
| APPENDIX: Quality Performance Assessment Framework | 27 |
| REFERENCES | 30 |
| ACKNOWLEDGEMENTS | 31 |

INTRODUCTION



Since the fall of 2008, the Center for Collaborative Education (CCE) has partnered with the Nellie Mae Education Foundation (NMEF) to explore the potential power of performance assessments to transform student learning and teacher practice. Through the Quality Performance Assessment Initiative (QPA), we have collaborated with a diverse group of 20 schools to strengthen and document local assessment systems and to implement common performance assessments across schools. The schools include urban, rural, and suburban schools, and represent charter, Pilot¹, and district schools across Massachusetts, New Hampshire, and Rhode Island.

QPA has developed the following definition of performance assessment:

Quality Performance Assessments are 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.

Quality Performance Assessments: Harnessing the Power of Teacher and Student Learning, and the accompanying *Quality Performance Assessment: A Guide for Practitioners*, describe specific actions and strategies for faculties as they undertake the challenge of developing a strong local assessment system with performance assessment as the cornerstone. The stories from the field share the experiences and entry points of three participating schools.

The Need for Quality Performance Assessments

The purpose of any assessment is to provide meaningful information about what a learner knows and can do. But what kind of information and how useful it is depend on the nature of the assessment. While current high stakes standardized tests provide information to state and district leaders to help them identify areas needing improvement and gaps in achievement of standards, they do not provide timely, useful feedback to teachers and students (Stiggins, 2008). Students and teachers need continuous focused data to know what

students are learning now and what they need to learn next.

Current standardized tests do not do a good job of assessing complex, multi-step intellectual tasks that are at the heart of the 21st century skills all students need. Teacher designed performance assessments that are linked to curriculum and instruction have the potential to provide the feedback that is so crucial to improving student learning and achievement.

When teachers—professional practitioners who are closest to the learners—are trained to be experts in developing and using assessments, we will see dramatic strides towards closing achievement gaps.

The recent adoption of the Common Core State Standards (CCSS) by 45 states has placed significant attention on the assessment of student performance at the national and state levels. As Secretary of Education Arne Duncan noted in a 2010 press conference, “bubble tests” have failed, and higher quality assessments are needed to “better measure the higher order thinking skills so vital to success in the global economy of the 21st century and the future of American prosperity” (Duncan, 2010).

ENGAGING IN MEANINGFUL WORK

Seniors at Fenway High School in Boston, Massachusetts, are expected to write and present a position paper that demonstrates their ability to think deeply about a subject and to write meaningfully. The Senior Position Paper is aligned with the Common Core State Standards for English Language Arts as well as with the Fenway Habits of Mind—perspective, evidence, connection, relevance, and supposition. To graduate, seniors must present and defend an idea, write persuasively, use appropriate voice, conduct relevant research, use appropriate citations, and revise their work. Common rubrics support this work and shape teacher practice and student expectations at each grade level by fostering common understanding. Grade-level expectations are further solidified through the creation of anchor papers used by teachers and students. (See “Story from the Field” on p.15)

¹ Pilot schools are unionized, district schools that have charter school-like autonomy over key areas: budget, curriculum and assessment, governance, professional development, school calendar and scheduling, and staffing.



Rigorous standards outlined in the Common Core State Standards set a new target ensuring that all students are ready for college and career. Considering that 40 percent of first year college students need to take at least one remedial course upon enrollment, much work needs to be done before this new target is met (Conley, 2007).² This statistic demonstrates a large gap between public K-12 school assessment of college readiness and higher education measures of actual readiness.

Regardless of the standards or the assessments that measure them, all students should have the opportunity to engage in meaningful work that prepares them for the 21st century workplace. The local level—within each school—is where we must ask: Is this assessment meaningful to students? Does this assessment measure real world skills and knowledge? Does this assessment provide feedback that motivates students to continue learning? The new standards and assessments create a seismic shift in the national assessment conversation. However, the corresponding shift in the achievement of our nation's students will occur only if we use new standards to rethink the intersections of teaching, learning, and assessment within each school and classroom and focus on deeper understanding. Practitioner-developed performance assessments must be a large part of the equation if assessment is to help all students in the United States achieve the prosperity of which Secretary Duncan spoke.

Policy experts Chester Finn and Michael Petrilli caution, “Standards describe the destination that schools and students are supposed to reach, but by themselves have little power to effect change. Much else needs to happen to successfully journey towards that destination” (Finn and Petrilli, 2010, p.2). If the destination is college and career readiness for every student, it is critical that schools prepare each student for this journey.

The reality of accountability for local districts and schools is clear. The simultaneous transition by 2014 to a set of new, more complex standards and a new, more complex test designed to measure those standards requires that districts, schools, teachers, and students begin preparing now. By developing and strengthening the local assessment systems today, schools and districts can create the foundational skills teachers and students will need in the future.



Quality Performance Assessments are 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.

² David Conley defines college readiness as “the level of preparation a student needs to enroll and succeed—without remediation—in a credit-bearing general education course at a postsecondary institution” (Conley, 2007).

A FRAMEWORK FOR HIGH QUALITY PERFORMANCE ASSESSMENT SYSTEMS



The challenging work of developing and learning to use high quality performance assessments to make decisions about curriculum and instruction cannot be done by teachers or schools working in isolation. CCE's Quality Performance Assessment (QPA) Initiative emphasizes the implementation of high quality common performance assessments³ across schools to drive professional development in assessment practice and improve teaching and learning of 21st century and higher order thinking skills.

In anticipation of the new assessments aligned to the CCSS, the QPA Initiative supports the incorporation of common performance assessments, aligned to English language arts and content literacy Common Core State Standards, into existing teacher practice. Our goal is to use the lessons from this first discipline to extend the model to mathematics, science, and social studies/history in the future.

Strengthening the assessment literacy of all educators improves every aspect of teaching and learning. Teachers possess assessment literacy when they understand “the difference between sound and unsound practices in assessment, evaluation, and communication” (Stiggins, 1999). Assessment experts from the Forum for Education and Democracy (Wood, Darling-Hammond, Neill, & Roschewski, 2007) note ongoing formative assessments, including performance assessments, can be “responsive to emerging student needs and enable fast and specific teacher response, something that standardized examinations with long lapses between administration and results cannot do.” Performance assessments can provide meaningful, real time information for students, teachers, parents, and administrators, and can be a springboard for improving teacher practice. They also note, “As teachers use and evaluate [performance assessment] tasks, they become more knowledgeable about the standards and how to teach to them, and about what their students’ learning needs are (Wood, et al. 2007).”

Student learning is also enhanced during performance assessment as students adjust their strategies and make timely corrections in response to targeted feedback from their instructors. This “assessment for learning,” differs from tradi-

tional assessments that function as a separate measurement of learning. Thus, local assessment systems that include performance assessment have the potential to improve both student learning and teacher performance. Further benefits of assessment systems with embedded performance assessment include greater teacher buy-in, increased teacher collaboration, and increased capacity to make mid-course corrections based on formative data (Wood, et al. 2007). When teachers are engaged as designers of performance assessments and skilled assessors of their students’ performance, the impact on curriculum and instruction can be profound.

By building school-wide assessment literacy and encouraging teachers and school leaders to take ownership of the local assessment system, the QPA Initiative lays the foundation for strong local assessment practice, creating a bridge to meaningful learning, college and career readiness, and success on the next generation of assessments.

Fundamental Elements of Quality Performance Assessment

The purpose of the QPA Framework⁴ is to provide guidance to teachers and administrators on how to design a performance assessment system. It describes a set of processes that can be implemented over time. Many aspects of the QPA Framework can be integrated into an existing system without a comprehensive overhaul. The following elements guide the work of creating quality performance assessments:

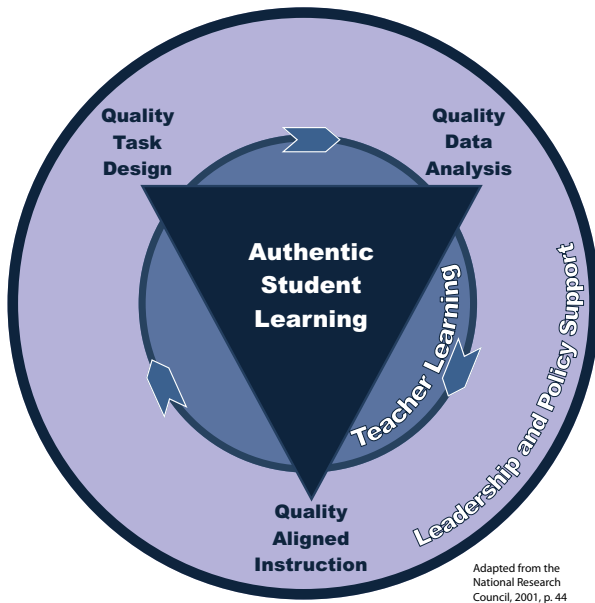
- **Aligned Instruction**—To gain knowledge and skills, all students need instruction, based on college and career readiness standards, that is accessible to their learning strengths and needs.
- **Task Design**—Valid assessment tasks include determining appropriate levels of content and cognitive complexity, setting clear criteria for success, and ensuring accessibility for all students.
- **Data Analysis**—By examining student work and score data, practitioners infer important information for planning future instruction and assessments.⁵

³ We define common performance assessments as those in which the same assessment is administered to all students in a grade and subject across multiple classrooms, schools, or districts.

⁴ Readers can access the complete version of the QPA Framework in the Appendix.

⁵ The criteria are aligned with the three vertices of the National Research Council Assessment Triangle: Cognition, Observation, and Interpretation (National Research Council, 2001)

QPA Framework



The graphic above illustrates how these components form a cycle with authentic student learning at the center. QPA defines authentic learning as learning that is meaningful to students, measures complex skills and content in multiple modes, and is transferable to new situations. It is critical that the standard for proficiency is clear to students and that clear feedback is provided to improve performance.

Essentials for Success

As schools undertake the important work of performance assessment, three Essentials must be in place to ensure successful design and implementation: technical quality, professional communities of practice, and leadership and community support (Tung and Stazesky, 2010).

TECHNICAL QUALITY

High quality performance assessments must have *technical quality*. They must be valid, reliable and sufficient. *Valid* means the assessment measures what it was intended to measure. *Reliable* means a group of teachers (or scorers) can agree on what a rating means and score it the same way. *Sufficient* means a student has been given a complete opportunity to demonstrate mastery. Assessments must be aligned to standards and designed to produce evidence of student learning. Without technical quality there is no guarantee an assessment system has evaluated student learning fairly and completely. When there is technical quality in assessment,

teachers become aware of whether or not they are in fact teaching what is being assessed, and can adjust their instruction accordingly.

PROFESSIONAL COMMUNITIES OF PRACTICE

Implementing an effective performance assessment system requires a cultural shift as well as a deepening of professional knowledge and skills. Teachers need to learn to use new tools to develop and implement performance assessments, and to score work together so they can internalize common expectations and score with consistency. There must be ongoing conversations focused on expectations, student potential, the role of effort in performance, and identifying teacher practices that need to change. This work takes time and trust in order to de-privatize teaching practice and encourage teachers to share teacher and student work. Because consistent scoring of performance assessment occurs in an environment where practitioners can question expectations for students and critique the instructional practice of their colleagues, de-privatization requires a collaborative school culture with a non-defensive faculty. District and school leaders must ensure there are ample resources and time to engage faculty in professional development and in building effective communities of practice.

Investing in the creation of teacher-led assessment systems is the only way to ensure that our students will learn the skills outlined by the CCSS. Such systems, reflective of teacher expertise, will ensure that curriculum, instruction, and assessment are closely aligned and result in authentic student learning.

The deep engagement of teachers with performance assessment must start in pre-service programs so that new teachers are prepared to implement meaningful assessments. Teacher education programs can create communities of practice where student teachers align, design, and analyze quality performance assessments before their first teaching assignment. Engaging in such rich tasks together has the potential to create the next generation of teachers who see collaboration around assessment and instruction as part of their professional responsibility.

COMMUNITIES OF PRACTICE IN ACTION

In the Pentucket Regional School District (Massachusetts) teachers prepare students for life beyond secondary school by teaching and assessing key Habits of Learning (HOL). The HOL include thinking and reflection, communication, collaboration, independence, and creative exploration. They are part of the Pentucket local assessment policy. In order to develop a common understanding of what teachers are expected to teach and students are expected to learn, Pentucket district leaders provide all teachers and students with detailed descriptions of the Habits of Learning. The district policy requires that HOL be taught explicitly and assessed throughout the school year from pre-kindergarten through high school. The HOL represent important local standards that are integral to graduation requirements and are meaningful indicators of student performance across grade and subject levels.

Pentucket Assistant Superintendent Dr. William Hart designed sessions in which teachers from the middle school came to the high school to share examples of HOL assessments they had developed. One participant noted, "This time [was] invaluable to share ideas, motivate teachers to improve their practice, and allow teachers time to continue to explore the use of performance assessments and how the HOL rubrics can work to improve our classroom instruction and outcomes." (See *Story from the Field* on p. 18)

LEADERSHIP AND POLICY SUPPORT

Successful efforts to adopt quality performance assessments have been driven by a high level of support from teachers, parents, community members, and school and district officials. The more all stakeholders participate in building the foundation of the performance assessment system, the more school leaders will be able draw upon this base of support in the future. The need for such political will makes it especially important to field test, fine tune, and scale up the system slowly, particularly if there are high-stakes outcomes based on student performance.



THE QUALITY PERFORMANCE ASSESSMENT FRAMEWORK IN ACTION— THREE STORIES FROM THE FIELD



The following stories examine the work of a charter school, a Pilot school and a school district. Each story highlights a particular aspect of the QPA Framework and concludes with a description of actions taken and recommended next steps. Together, they illustrate multiple possible entry points for engaging in this work.

STORY 1

Quality Data Analysis: Cape Cod Lighthouse Charter School Assessment Validation Sessions

Cape Cod Lighthouse Charter School chose to focus on technical quality as an entry into performance assessment. The faculty wanted to ensure that the assessment tasks they developed were *valid*—producing intended information about student learning. The faculty explored the following key questions:

- Does the assessment provide the information about mastery of standards/content for which it was designed?
- Do student work samples demonstrate proficiency for the subject and grade level?
- Do teachers and other school faculty use data from performance assessments to inform curriculum planning, instruction, and (re)design of assessments?

An important step in determining validity is ensuring that learning assessments are clearly aligned to standards and that they measure student performance on the intended standards. In order to meet validity requirements, assessments must be appropriate for the standards being measured. For example, a multiple-choice test would not be a valid measure of a student's ability to write a cohesive, well-organized argument, nor would it measure his or her ability to express and defend ideas orally.

THE QPA FRAMEWORK

Quality Data Analysis

Quality assessment data analysis involves examining both student work and score data for *technical quality*. Assessments must be valid, reliable and provide sufficient evidence of learning. *Valid* means the assessment measures what it was intended to measure (both content and intended level of rigor). *Reliable* means a group of teachers (or scorers) agree on what a rating means and can score it the same way. Reliability is essential because assessment data leads to high and low stakes actions and decisions. To ensure that all students are demonstrating mastery sub-group performance should be examined for bias in score results. *Sufficient* evidence means students have been given a complete opportunity to demonstrate mastery resulting in enough evidence of learning being collected. Without technical quality there is no guarantee that an assessment system has evaluated student learning fairly and completely. Conclusions from the data analysis provide information to practitioners for planning future instruction and assessment.



Working with professional development support from CCE, Cape Cod Lighthouse Charter School (CCLCS) teachers used a structure for critical review to share and critique assessment tasks: the Assessment Validation Protocol.⁶ Several teachers prepared and presented an assessment they wished to validate. The assessments ranged from an independent reading project to a foreign language assessment to a social studies museum artifact project. In preparation for the validation session, teachers gathered all documents related to their assessments, including prompts, standards maps, rubrics, and scaffolding materials to share with their colleagues. In addition, they selected samples of proficient student work that would serve as evidence of students' ability to demonstrate mastery of the selected standards on the assessment.

Steps in the Assessment Validation Protocol

1. QUALITY ALIGNED INSTRUCTION

The session begins with a cross-disciplinary group of four to six teachers—the validation team—reviewing all documents including an “Assessment Validation Cover Sheet” that lists the standards to be assessed. First, the team spends 10 minutes ensuring that the assessment is aligned to standards. To do this, the team carefully reviews the standards being assessed by the given task and compares them to the submitting teacher’s expectations. In addition, the team ensures these standards assess students on the competency level of the standard referred to as “depth of knowledge” (Webb, 1997) levels. Assessments measure skills ranging from basic (e.g., recall and memorization) to complex (e.g., critiquing and presenting multiple viewpoints). To validate alignment, the validation team must indicate whether or not it believes each standard has been accurately assessed, discuss findings, and reach 80 percent consensus.

2. QUALITY TASK DESIGN

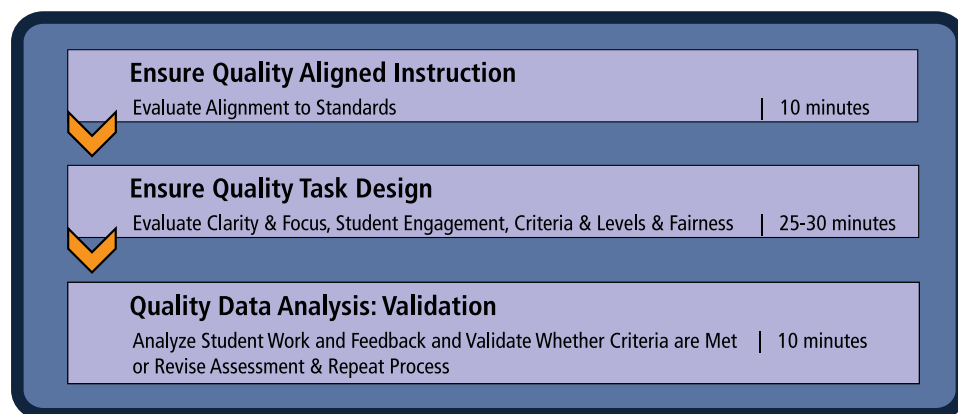
Next, the team spends 25-30 minutes evaluating the assessment’s clarity and focus, the opportunities provided for student engagement, the appropriate use of rubrics or scoring guides, and the degree of fairness. The validation team looks for evidence the assessment is linked to instruction, requires students to actively engage in the task, is clear and easy to understand, and uses an appropriate rubric for scoring. To evaluate fairness, team members also determine whether the task is free from stereotypes, whether it is fair to students of all races, cultures, ethnicities, religions, etc., and whether it uses the principles of “universal design” (Rose and Gravel, 2010)—that is, whether the task uses language and a format that all students can understand. Again, in order to validate this aspect of the assessment, the team must indicate whether or not it has found evidence of each criterion, discuss its findings, and come to an 80 percent consensus.

3. QUALITY DATA ANALYSIS: VALIDATION

Finally, the validation team determines whether or not the criteria for validation have been met by analyzing the student work and all feedback from the alignment and design sections of the protocol. If the team is able to reach 80 percent consensus that there is evidence for each assessment criterion in the validation protocol, the task is validated. If the team fails to reach consensus, or if it determines that any criterion was not met, it provides feedback for revision. A member of the validation team meets with the submitting teacher, reviews the feedback, and makes a plan for resubmission if necessary. The process can be repeated until validation is achieved.

⁶ The QPA assessment validation protocol is adapted from the work of Karin Hess (Hess, 2009).

Figure 2: QPA Assessment Validation Protocol



CRITERIA FOR A VALID ASSESSMENT

Assessment Is Aligned

- Assessment is aligned to specific content standards
- Assessment is at the appropriate depth of knowledge to assess the standard
- Assessment is aligned to what is intended to be assessed and will elicit what students know and can do related to chosen standards
- Assessment is scheduled to provide enough teaching time to allow students to succeed

Assessment Has Clarity and Focus

- Assessment addresses an essential issue, big idea, or key concept or skill of the unit/course
- Assessment is linked to ongoing instruction (within a unit of study/course)
- Clear directions indicate what the student is being asked to do
- Assessment includes what will be assessed individually by the student (even if it is a group task)

Assessment Allows for Student Engagement

- Assessment provides for ownership and decision-making and requires the student to be actively engaged
- Assessment provides authenticity and reflects a real world situation or application

Assessment Uses Appropriate Criteria and Levels

- Rubric(s) or scoring guide(s) assess all intended parts of content standards
- Exemplars/anchor papers illustrate expectations aligned to standards

Assessment Is Fair and Unbiased

- Material is familiar to students from different cultural, gender, linguistic, and other groups
- Task is free of stereotypes
- Students have equal access to all resources (e.g., Internet, calculators, spell-check, etc.)
- Assessment conditions are the same for all students
- Task can be reasonably completed under specified conditions
- Rubric or scoring guide is clear

Assessment Adheres to the Principles of Universal Design

- Instructions are free of wordiness or irrelevant information
- Instructions are free of unusual words students may not understand
- Format/layout conveys focus of expected tasks and products
- Format clearly indicates what actual questions or prompts are
- Questions are marked with graphic cues (bullets, numbers, etc.)
- Format is consistent

Assessment allows for Accommodations for Students with IEPs/504 Plans

The Validation Process

CCLCS started the validation process in January 2011. Teams met three times before the end of the school year and reviewed 12 assessments that included tasks in every subject and all three grade levels. These assessments were reviewed by interdisciplinary validation teams, but not a single one was validated on the first round.⁷ Why was this lack of validated assessments considered a success? A lack of initial validation meant the process was successful in uncovering the assessment creators' blind spots and assumptions, so that the assessments could be refined for future use. For example, one question in the validation protocol asked if the scoring guide was clear. Of the 12 assessments presented at CCLCS, only three validation teams felt the rubric was clear on the first review. Presenting teachers could then take their colleagues' precise feedback focused on the clarity of the rubric/scoring guide and revise before giving it to students again. Once revisions were made, teachers could resubmit to the validation team. CCLCS resumed this process in the fall of 2011 to validate the first 12 assessments.

Feedback from the validation sessions included the following comments:

- **Fairness**—*Assessment is unfair because the lack of clarity and specifics in the project guidelines means that an “A” student will get it, but a struggling student will require more guidance to be successful.*
- **Clarity and Focus**—*More detail about the process and intent would be beneficial to students. How do students know what not to do to get a perfect score?*
- **Student Engagement**—*Structure a time and protocol for students to compare cars to see why one performed better than the other based on the laws of motion.*
- **Criteria and Levels**—*The rubric needs work. It needs to be easier to read. It is missing a few categories such as display and presentation, and quality of writing.*
- **Alignment**—*Not only ask, ‘What changes did you make?’ but ask, ‘Why are you making them and how are they related to the laws of motion?’ Each student learns about one system in depth, but the standard calls for them to master all the body systems. Can you create a test or way they learn from other students to ensure they understand all systems?*

One teacher commented that the process is helpful because it supports teachers in “getting at the essence of where problems lie in our assessments and tweaking them so the quality of the assessment is improved.” CCLCS has created a community of practice where teachers collaborate and provide valuable feedback to each other to improve the validity of their assessments.

Lessons Learned—How the Cape Cod Lighthouse Charter School Assessment Validation Sessions Reflect Best Practice

Technical quality, one of the three Essentials for a performance assessment system, is at the heart of the validation sessions. Rather than relying on basic intuition or chance to ensure that tasks are valid, CCLCS teachers decided to systematically determine whether assessments met the validation criteria. By doing so, CCLCS is well on its way to making sure measures of student learning provide relevant, meaningful information about what students know and can do to students, parents, teachers, and local administrators. In addition, the process allows teachers to ensure tasks are aligned with standards and teaching and that they are fair.

Highly motivated Cape Cod Lighthouse Charter School teachers who support performance assessments have driven technical quality by engaging in validation sessions. By investing additional time and effort to develop assessment literacy, these teachers have created their own rewarding style of practice.

One validation session participant noted its potential to transform practice, stating, “Looking at assessments with a critical eye was extremely beneficial and will not only help me become a better teacher, but will also certainly enhance my students’ learning and improve their depth of knowledge.” This kind of feedback can have a positive impact on practice throughout the school. Teachers who support performance assessments may encourage their peers to follow suit. As teachers and leaders build fluency with performance assessments, they also build their school’s or district’s capacity to develop and implement professional development activities that facilitate this work.

⁷ Throughout this paper, we refer to unpublished documentation and artifacts shared with QPA by the schools, including validation feedback, PD evaluations, teacher reflections, personal communications, and teacher and student work. Because they are unpublished, they do not appear in the References section. For further information about these types of documentation, please contact QPA directly.



Looking at assessments with a critical eye was extremely beneficial and will not only help me become a better teacher, but will also certainly enhance my students' learning and improve their depth of knowledge.

—CCLCS Teacher

The CCLCS Assessment Validation Sessions also reflect important aspects of the QPA Framework. This practice demonstrates how a school or district can

- **Analyze assessments for alignment to prioritized standards**—Aligning standards and assessments does more than just ensure the student work and data provided have a clear purpose and use. It also ensures that all students have had the opportunity to learn the standards that are measured. By validating their assessments, teachers at CCLCS are taking steps to provide a rigorous, equitable education to all students. This process also has the potential to improve practice as teachers become aware of whether or not they are in fact teaching what is being assessed and adjust instruction accordingly. Without this important collaborative critique process, assessments run the risk of being irrelevant to both students and teachers, because they may not provide appropriate information about what students know and can do.
- **Conduct meaningful cross-disciplinary conversations**—Another important aspect of the validation sessions is that they are cross-disciplinary. As groups of teachers work across subject areas, they have important conversations about expectations for performance across subject areas. This discussion provides an opportunity for teachers to not only align standards and assessments, but also to align their expectations. While traditional policies may leave students confused about what is being asked of them from one course to the next, teachers, working collaboratively demystifies expectations for students, allowing them to anticipate the level of work being demanded in all courses.

Possible Next Steps

As teachers and school leaders engage in this work, it is important to remember that it is an iterative process.⁸ A possible next Step in the design and implementation of a local performance assessment system for CCLCS may be: *Step 7: Determine whether outcomes on teacher-created performance assessments, and the interpretations made about learning, are closely related to students' outcomes on other measures of the same standards.*

⁸ To support schools in planning their next steps, QPA has developed the “Steps to Quality Performance Assessment” (Read all 10 Steps on page 26), aligned to the QPA Framework. Each story from the field concludes with an analysis of the Steps schools have taken from this list.

CCLCS's local assessment system includes a series of benchmark performance assessments students must pass in each grade level. Reviewing student scores on these benchmark assessments and comparing them to how students perform on other indicators—including grades, standardized tests, and diagnostic tests such as reading assessments—provide important information about students' learning needs and about the assessments themselves. Student scores on local performance assessments should be related to scores on other measures of the same standard. Perfect alignment of performance assessments and other evaluations of students' competencies should not be expected, as performance assessments tend to measure a higher level of cognitive complexity and to assess multiple standards simultaneously.

However, because student achievement cannot be understood by relying solely on standardized test information performance assessment data is essential for understanding student learning. Analyzing local and state assessment data creates a fuller picture of students and the work of the school.

In addition, QPA Step 8 suggests that schools and districts: *Collect evidence to document consistency in scoring and calculate a reliability score for each important assessment.*

Assessment data is useful if it is consistent and reliable. Scoring consistency implies that information gathered from one measure of student learning does not vary significantly from teacher to teacher, or in time. Scores generated by one teacher must be the same as those generated by another. Since some level of subjectivity is always involved in scoring with rubrics, a small amount of variation should be expected.

In order to achieve scoring consistency, time and resources must be invested in scoring sessions. Similar to the validation sessions, scoring sessions involve teachers working in groups to score student work, using rubrics, and coming to a consensus about scores. Extensive analysis and refinement of rubrics, ongoing scoring practice, and deep conversations about consistency can help teachers attain a high level of agreement about proficient work. While this work requires a significant investment of time, it is critical to ensuring that assessment data is meaningful and of high quality.



Student achievement cannot be understood by relying solely on standardized test information. Performance assessment data is essential for understanding student learning.

STORY 2

Quality Task Design: Fenway High School Junior Review, Senior Institute, and Senior Position Paper

Fenway High School in Boston, Massachusetts, engaged in the performance assessment design process. Teachers and school leaders wanted to ensure that students demonstrated mastery of college and career ready skills through a series of benchmark assessments completed in their junior and senior years. This Fenway High School (FHS) story explores the following questions:

- What are the criteria being used to assess mastery?
- Are there samples of student work to illustrate work at each performance level?
- Do students at all levels have sufficient opportunity to demonstrate mastery through the assessment?
- Is the assessment appropriate in content and skill level for the grade being assessed?

Performance Assessment for Authentic Learning

Effective assessment development begins with clarity about what students at each grade level should know and be able to do. A common understanding among faculty about content and cognitive complexity in the grades they teach and adjacent grades guides the design of prompts and scoring tools. Documentation of the assessment design and a validation process build awareness of expectations, allowing appropriate performance levels to be set at each grade level.

At FHS, assessments are used to measure authentic learning in ways that give all students the opportunity to demonstrate mastery of key standards as they progress through the grade levels. FHS's assessment policy has resulted in a series of aligned benchmark assessments that are part of the Junior Review and Senior Institute. Underlying this policy is the belief that measures of learning should “be used for the benefit of the child, be intellectually and culturally fair to all students, support individual development not competition for grades, [and] reflect the use good Habits of Mind.”⁹ As part of this policy, “Fenway actively explores better ways to determine when...students are prepared to do the kinds of work and study required in the real world,” including through “classroom-based diagnostics, portfolios, project and assignment outcomes, exhibitions, and tests.”

⁹ Habits of Mind are attributes or behaviors that students display when they are using their minds effectively and efficiently (Costa, 2008).

THE QPA FRAMEWORK

Quality Task Design

Effective assessment development begins with clarity about what students at each grade level should know and be able to do. A common understanding among faculty about content and cognitive complexity in the grades they teach and adjacent grades guides the design of prompts and scoring tools. Documentation of the assessment design and a validation process build awareness of expectations, allowing appropriate performance levels to be set at each grade level.





Fenway actively explores better ways to determine whether our students are prepared to do the kinds of work and study required in the real world

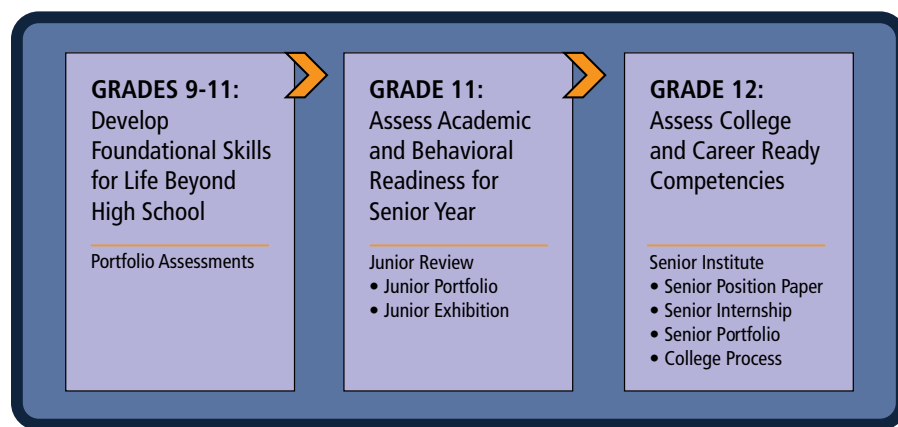
—Fenway Handbook

WHAT IT TAKES TO BE A FENWAY GRADUATE

- **Senior Internship:** All Fenway seniors are required to fulfill a full-time six-week pre-professional internship during the last term of Senior Institute. The internship deepens the student's understanding of a field of work and the role education can play in preparing the student for success. A portfolio that documents the experience, including a research component, is part of the requirement.
- **Standardized tests:** A variety of standardized tests are offered to students in their senior year. Some are required and others are optional.
- **Planning for the Future:** More than 90 percent of Fenway's students go on to college. An important part of senior year is spent thoroughly researching post high school options and keeping track of college application deadlines.
- **Course work in math, science, and humanities:** Courses can be either at the high school or college level. Additional courses may also be included.
- **Senior Portfolios in the major subject areas:** math, science, and humanities
- **Senior Position Paper**
- **Science Fair**

—Fenway High School Student Handbook

Figure 3: Fenway High School Benchmark Assessments



All eleventh graders at FHS are required to complete the Junior Review, an exhibition of learning captured through portfolios. According to the Fenway High School Handbook, “Portfolios are purposeful collections of student work that show understanding of different and important topics and skills.” Juniors at FHS choose a series of assessments that demonstrate their mastery of benchmark standards to include in their portfolios to prove they are prepared for senior year. During the Junior Review, students present a range of academic content to “prove their ability in the major content areas, as well as demonstrate intellectual and social readiness for the challenging senior year.” This presentation represents vertical alignment, as it serves both as an eleventh grade exit assessment and as a foundation for students’ senior year graduation requirements.

Students who successfully demonstrate readiness for senior year through their Junior Review are inducted into the Senior Institute. The Fenway High School Handbook notes, “The Senior Institute captures and crystallizes the skills and intellectual maturity learned during the previous three years at Fenway.” During Senior Institute, students are again assessed on both academic and non-academic criteria to determine their readiness for the next step—college and careers. The Institute culminates in a series of assessments through which students have the opportunity to demonstrate their competency in key graduation requirements. Beyond traditional measures such as standardized tests, students show that they are able to meet more authentic demands through a range of performance assessments which include presenting a portfolio of work, completing a six-week internship, completing the college application process, and completing the Senior Position Paper.

While students at FHS have noted that meeting senior year requirements can be challenging, one FHS graduate notes, “Here at Fenway there is a lot you have to do to get out of high school. At least you know that if you are graduating from

this school that you are not getting off easy. You have proven you have met the tough requirements” (Stefanakis, 1997).

FHS’s thoughtfully developed local performance assessment system helps support student learning. Beginning freshman year, teachers and staff at FHS carefully scaffold the knowledge and skills students must master by building authentic assessments into their curricula. By the time students reach senior year, they have worked on their ability to think critically and reflectively about their learning for three years. Further, as the Fenway Handbook notes, “Students edit and improve [assessment] pieces and save them in portfolios where they are available as evidence and for reflection and use in Junior Review or for Senior Institute Portfolios.” By senior year, students have thus had ample opportunity to not only show what they know and can do, but to reflect on their own learning in order to improve.

Quality Task Design in the Senior Position Paper

Seniors at FHS are expected to write and present a Position Paper that demonstrates their ability to think deeply about a subject and to write meaningfully. The Senior Position Paper is aligned to the Common Core State Standards (CCSS) for English Language Arts as well as the Fenway Habits of Mind—perspective, evidence, connection, relevance and supposition. More specifically, “the goal of the paper is to focus on developing the skills of strong writing, critical thinking, and presenting an argument that are necessary for success in college.” It calls for careful thought and effort, as students are assessed both on the quality of their ideas and the quality of their writing. In order to graduate, seniors must present and defend an idea, write persuasively, use appropriate voice, conduct relevant research, use appropriate citations, and revise and edit their work.

A CLOSER LOOK AT THE FENWAY HIGH SCHOOL SENIOR POSITION PAPER

Write a final Senior Position Paper related to a person or event. A senior must independently write, revise, and edit a paper arguing a position that is related to his/her person or event. A senior must be able to use his/her own voice to persuasively argue that position. The paper must show that a senior can: identify a key issue and a personal position, gather information from at least three sources, logically argue for the position, use supporting evidence from three sources, and acknowledge opposing perspectives. A senior's Position Paper must have appropriate format, clear structure, paragraph unity, and coherence; and correct conventions for citations and bibliography, and for spelling, grammar, and mechanics. As in other writing pieces, a senior must show evidence of having followed a writing process and of revising and editing independently. All note cards, outlines, and drafts must be filed with the final Position Paper.

— *Fenway High School Senior Portfolio Guidelines*

Teachers have developed a common project sheet and rubrics for the Senior Position Paper that shape teacher practice and student expectations by fostering common understandings. Grade-level expectations are solidified through the creation of anchor papers. As part of a local assessment portfolio, teachers at FHS documented evidence of the performance level for Senior Position Papers for each level of the rubric. These anchor papers serve as models for students, so that they have a clear understanding of their target and can demonstrate mastery and meet the graduation benchmark.

The Senior Position Paper proficiency requirement means that students revise their Senior Position Paper until it reflects a standard of quality that merits graduation. FHS teachers frequently review the standard to ensure that it is set appropriately and aligned to the level required for college readiness. In 2010, as teachers worked together in a community of practice to review anchor papers and submit documentation for FHS's Local Assessment System Portfolio, FHS teachers decided to increase the level of proficiency required on the Senior Position Paper. While the foundation for this expectation has been set, teachers at FHS will also need to align their instruction to achieve the new standard for this assessment and to “vertically align the quality of the Senior Position Papers...to ensure that the quality of the work is at the level required for college success.” Teachers can achieve this level of Quality Task Design only when they set consistent expectations collaboratively.

Lessons Learned—How the Fenway High School Benchmark Assessment Process Reflects Best Practice

The Junior Review and Senior Institute are the result of many years of the school leaders' clear vision and support for developing quality performance assessment tasks that measure authentic student learning at the appropriate grade level. As a member of the Coalition of Essential Schools and the Boston Pilot Schools—two networks that promote progressive practices to benefit all students—FHS has had ample political support to develop this policy.

Larry Myatt, founder and former co-director of Fenway High School, describes this vision:

We all dreamed that if we put in place the right people with the right ideas during Senior year, we could stop them [students] from scraping by with a D minus. So our goal was to link work from internships, from a senior project, and from courses. Senior Institute was a way to customize these experiences and offer enrichment, remediation or acceleration. The notion was

to develop students' capacity for independent work, for making connections (Stefanakis, 1997).

This story of the development of the FHS assessment policy addresses the QPA Framework element of Leadership and Policy Support and Step 9 which suggests: *Document and adopt the assessment policy through a process that builds political will and support of all stakeholders.*¹⁰ Since an authentic performance assessment system requires significant time and effort on the part of students, teachers, and school leaders, buy-in from all parties is crucial.

Students and their parents must value creating portfolios, conducting presentations of learning, and revising student work multiple times. FHS demonstrates that this work can be explicitly linked to college or career ready outcomes, so that it is clearly viewed as a worthy investment of students' time and energy as well as of parents' support.

The FHS story also addresses the QPA Framework element of Teacher Learning in Communities of Practice and Step 10 which suggests: *Design professional development in communities of practice that supports all teachers in effectively implementing the policy.* Teachers must view these assessments as meaningful learning opportunities from which they can gather relevant information about what students know and can do. Such a system demands more of teachers, as they will spend more time and thought developing assessments and providing ongoing feedback to students. Therefore, time and resources for professional development for this work are imperative. School and district leaders must spearhead this work, providing the vision and support as teachers continue along a challenging but worthwhile assessment path.

Beyond these policy lessons, the FHS assessments are models of important aspects of the QPA Framework element Quality Task Design. Specifically, they demonstrate the Step 3: *Ensure that each assessment's content and complexity are appropriate for the assessment grade level, based on the school's established content sequence and grade-level standards.*

Vertical alignment, the process of creating a cohesive scope and sequence of standards across grade levels, has been an important driver of the development of the Junior Review and Senior Institute. This process ensures that by junior year, students have been prepared, through relevant curricula and meaningful measures of learning, to meet the demands of senior year. Since FHS seeks to foster high-level competencies and the practical skills needed to navigate and succeed in the postsecondary world, these competencies have been built into the academic scope and sequence. Beginning in ninth grade, FHS students are taught to think critically and reflectively about their learning, make informed decisions

¹⁰ For full list of QPA Steps to Quality Performance Assessment see page 26.



Here at Fenway there is a lot you have to get out of high school. At least you know that if you are graduating from this school that you are not getting off easy. You have proved you have met the tough requirements.

—Fenway High School Student

about their academic progress, self-assess their mastery of important standards, write authentically and for meaning, revise their work, and communicate effectively about their learning to diverse audiences.

In addition to scaffolding the content and skills needed to be successful beyond high school, FHS' assessment policy creates clear benchmarks that distinguish the cognitive complexity required from one year to the next. While Junior Review requires students to present portfolios and to reflect on their academic learning and non-academic habits, the Senior Institute requires a different set of skills. As seniors leave high school and enter college or the workforce, they must be prepared to meet the demands of these environments. Senior Institute prepares them by including an internship component as well as by requiring them to conduct relevant research on the college application process. These tasks serve as assessments of skill sets that extend beyond basic junior-year requirements and exemplify the sort of cognitive demands students will encounter beyond high school. FHS's local assessment system reflects the standard of readiness for college and career that the Common Core State Standards demand of all schools.

FHS's policy also addresses Step 5 in the QPA Framework: *Provide a sample of student work at each performance level that illustrates work at that level for each assessment.* Anchor papers help ensure that both teachers and students have a clear understanding of performance standards. FHS teachers identify anchor papers at each performance level that model expectations for Senior Position Papers. These anchor papers, along with rubrics, are provided to students as they complete their own Position Papers in order to facilitate their progress towards this important benchmark. Such a practice is integral to a high quality task design, because

students have more opportunities to attain their goals when they know what the end product should look like. Showing students models of strong work can also encourage them and their teachers to push themselves to reach the highest possible standard rather than settle for basic competency.

Fenway High School is well along in the process of ensuring high quality task designs within its local assessment system. Assessments are vertically aligned and measure standards that are clearly outlined, well documented at each grade, and linked to important, authentic outcomes. Differentiated models of student work serve as concrete guides for teachers and students as they strive for mastery.



STORY 3

Quality Aligned Instruction: Pentucket Regional School District Habits of Learning

The Pentucket Regional School District (PRSD) in Massachusetts designed an assessment policy that aligned instruction around Habits of Learning (HOL). The district leaders wanted to ensure that instruction and assessment were aligned to the skills students would need for success in college and career. The PRSD story explores the following questions:

- What does quality teaching look like within the context of a local performance system?
- How are local performance assessments aligned to national, state, district, and school standards; and how are these standards used to inform instruction?
- What scaffolding and support strategies can be used at the classroom level in order to support student mastery of standards?
- How do leaders design and implement policies and professional development to support performance assessment?

Teaching and Assessing Habits of Learning

All students need instruction based on college and career readiness standards. Once students have the opportunity to master these standards, assessments allow them to demonstrate what they know and are able to do. Further, information about how students approach learning can be as valuable to teachers as information about the actual content and skills they learn. College and career ready students employ a number of learning strategies to get results.

In PRSD, teachers prepare students for life beyond secondary school by teaching and assessing key Habits of Learning. The five HOL are thinking, communication, collaboration, independence, and creative exploration. They are an integral part of PRSD's local assessment policy. The district has determined that these skills represent “essential learnings” which are the critical skills, knowledge, and dispositions that are prioritized as being essential for every student to master (Dufour, Dufour, Eaker & Many, 2006). PRSD provides all teachers and students with detailed descriptions of the HOL in order to develop a common understanding of what they are expected to teach and learn. The district policy requires that HOL be taught explicitly and assessed throughout the

THE QPA FRAMEWORK

Quality Aligned Instruction

Teaching and assessment practices are interwoven. Instruction and performance assessments are aligned and based on essential learnings and appropriate national, state, district, and school standards. In addition, all students need instruction based on college and career ready standards. Once students have the opportunity to master these standards, assessments allow them to demonstrate what they know and are able to do.



school year from pre-kindergarten through high school. The HOL are powerful because they are integrated with content, not taught in isolation, resulting in deeper understanding and transfer to other contexts.

PRSD leaders documented the HOL policy in the student handbook. The policy states that students will be regularly assessed on the HOL using the PRSD HOL rubrics. District-wide performance assessments of the HOL in grades four, six, eight, and eleven took place for the first time in the 2010-2011 school year. Students were prompted to reflect upon how they used HOL in their courses throughout the school year and to prepare portfolios of work samples that demonstrated how they had used the HOL. Student presentations varied by grade level according to length, audience size, and preparation. For example, at the beginning of the semester, eleventh graders at Pentucket Regional High School were grouped into small cohorts that met monthly to self-assess their progress on HOL. Each cohort was assigned two faculty advisors to guide them through the work of preparing presentations.

In advance of the presentations, teachers and district leaders worked to develop common rubrics to assess students' demonstrations of HOL. These rubrics were made available to teachers and students before the district-wide presentations were required. Teachers worked to ensure that all students had clear understandings of the requirements and were given the opportunity to learn the standards before being assessed. At the end of the semester, juniors presented their portfolios to a panel of parents, teachers, administrators, and community members. These 20-minute capstone presentations have become part of PRSD's graduation requirements.

To support this work, administrators at the high school designated early release days for professional development to support the implementation of the HOL performance assessment system. For example, during one session middle school teachers went to the high school to share examples of the HOL assessments they had developed. "This time [was] invaluable to share ideas, motivate teachers to improve their practice, and allow teachers time to continue to explore the use of performance assessments and how the HOL rubrics can work to improve our classroom instruction and outcomes."

REFLECTION ON THE "HABIT OF THINKING"

It's one of those habits that is misunderstood because people think, okay, well I think every day. But it's a type of thinking that most people don't do every day—a reflective kind of thinking. So for the history book review the first section was to give what you thought of the book. I had to think back to my overall feelings when I read the book, not just say, 'Oh that was a good book.' Why did I think it was a good book? What parts of it made it a good book?

The next part of the book review was to say what you thought the author's purpose in writing the book was. And that to me was the biggest challenge—to think about what they were thinking when they wrote the book. I was really pleased with what I did because I realized that there is a whole different part of thinking—deep thinking.

—*Pentucket High School Student*

Towards a Cycle of Assessment and Instruction

As a long term goal, the district aims to have all teachers embed the HOL in instruction and use formative assessments to inform practice. One teacher remarked that using the Habits of Learning during instruction helped students improve their performance, thus demonstrating the power of performance assessment to function as assessment for learning.

Once students were introduced to the project, they were asked to consider how they could improve their creative exploration habits through the...assignment. Midway through, they again used the rubric to reflect on their progress. Finally, at the end, as part of their reflection, they evaluated themselves according to the rubric. The teachers were pleased with the result as it was clear that the rubrics were helpful in directing students to push their thinking.

—PRHS teacher

In his HOL presentation, a junior at the high school also revealed how the HOL of Communication can be embedded in instruction, further exemplifying the district's goal.

I think there's a big misconception with this habit. A lot of people when they think of communication they think of speaking and presenting, but communication is bigger than that. It's listening, it's observing and it's becoming better through observing other people present. It means successfully listening, speaking, observing, presenting, expressing, and delivering.

[In history class] We had seven groups to observe and to watch before us. By taking notes and carefully observing, we got to see what worked for them, what didn't work for them. Ultimately, when we sat down, we said, well this group, they did a lot of slideshow and a lot of fact-based information and that didn't engage the audience at all. And we wanted to do more than portray the facts of what our project was trying to tell them. We wanted to engage the audience. We made a slide

show and we posed questions to the audience throughout the entire slide show.

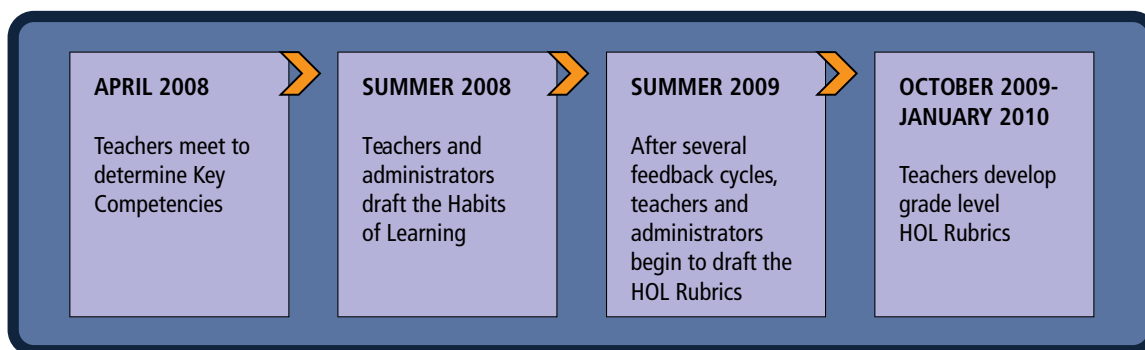
—PRHS Junior

Similar to the teacher reflection on the process of helping students develop creative exploration abilities, the student example demonstrates teachers ensuring that assessments are interwoven with teaching—that is, instruction is aligned to standards and provides all students the opportunity to master these standards. Teachers who embed assessments in instruction make sure they provide multiple opportunities for learners to engage with the standards prior to being assessed; students use rubrics to reflect on their progress and observe and critique their peers' work. District leaders at PRSD wish to foster this sort of teaching across subject areas, and grade levels.

Taking a Systemic Approach

Developing the HOL assessment system was a detailed process that evolved over time and continues to progress. PRSD assistant superintendent and architect of the HOL system, Dr. William Hart, notes that the development of the HOL rubrics and performance criteria reflects “months of exceptional work from teachers representing every school and every grade in the district,” and has been a “deliberate, thoughtful, and inclusive process.” First, teachers met in April 2008 to determine the key competencies students need to master in order to succeed in college. Next, teachers and administrators devoted several days over the summer to draft the HOL. After teachers in each school provided feedback on this draft, district administrators revised the HOL. In the summer of 2009, more than one year after the initial meeting, teachers and administrators at the Teacher Leader Summer Institute developed an HOL rubric. PRSD teachers then began the three-month process of creating grade-level HOL rubrics.

Figure 4: PRSD HOL Timeline



This process required not only time, but also a high level of community buy-in and committed leadership. At the district level, Dr. Hart and the District Professional Development Leadership Council supported this work. In order to fulfill the vision of creating “passionate learners prepared for an ever changing world,” the school district has made its mission to “regularly collaborate to balance content knowledge with HOL in a powerful learning environment...[and to] measure progress by locally created benchmark assessments and standardized tests” (Pentucket Regional School District, 2010). At the school level, PRSD principals have provided teachers with the time and resources needed for professional development in assessment literacy. According to Dr. Hart, this support includes an HOL Demonstration of Mastery Task Force to guide the process at each school. There is also support for students at each building level to guide them through their presentation and portfolio process.

Parents have been crucial to the success of the HOL vision. Parents attend the HOL presentations allowing them to witness the power of the process by being part of it. The following excerpts from an article in the Newburyport News, PRSD’s local newspaper, demonstrates this: “Karyn Williams (PRSD student) said the hardest part of the portfolio assessment was speaking in front of people for that length of time. ‘Afterwards my mom kept saying how proud she was of me,’ said Williams. Shannon Beaton (PRSD student) said the experience tapped strengths that aren’t always recognized through traditional assessment methods (Solis, 2011).” When parents become engaged in the process, they become advocates for performance assessment.

Lessons Learned—How the Pentucket Regional School District’s Habits of Learning Reflect Best Practice

Although PRSD’s work has come a long way, it is not complete. Dr. Hart concludes, “It will take several years to fine-tune our curriculum, instruction, assessment, and public demonstration procedures before we are at a point where we may be satisfied.”

In addition to highlighting important policy lessons—namely the importance of time, political buy-in, and professional development for local performance assessments—this story of the PRSD HOL addresses the QPA Framework element of Quality Aligned Instruction and Step 1 which suggests: *Determine graduation and promotion requirements, essential learnings, and/or habits of mind and work that focus the school on the most important standards for their students.*¹¹

The amount of time, thought, and effort the district has put into developing the HOL rubrics and assessments are a testament to the importance the district and schools assign to these essential learnings. The assessments provide meaningful, relevant information to students and parents, teachers, school leaders, and district policymakers—information that may indicate whether or not students are prepared to meet college-level expectations.

As one PRSD teacher noted, “The HOL performance assessment has the potential to be an example of how public schools can systematically assess students’ learning.” PRSD’s use of the HOL demonstrates quality aligned instruction in several ways. First, the HOL represent important local standards integral to graduation requirements. They are meaningful indicators of student performance that can be used across grades and subjects as part of a uniform assessment policy.

Second, the HOL are aligned to instruction and help drive teaching and learning. Integrated with content, HOL are not taught in isolation. Rather, they are used as a means of ensuring deeper learning and transfer of knowledge. Use of HOL assessments exemplifies how standards should be measured using end-of-year assessments, as well as formatively assessed through self-scoring and reflection.

¹¹ For full list of QPA Steps to Quality Performance Assessment see page 26.

Finally, PRSD has demonstrated the QPA Framework element of Leadership and Policy Support and Step 9 which suggests: *Document and adopt the local assessment policy through a process that builds political will and support of all stakeholders.* The Habits of Learning and performance assessments are a well-documented, integral part of the PRSD local assessment policy. District administrators have been intentional in their documentation of the HOL assessments, and have published relevant policies in the district handbook and on the district website. A commitment to parent communication enables parents to be informed participants in HOL presentations.

The project (HOL presentation) allowed him to articulate what was meaningful to him as a learner. It was one of those light bulb moments we parents love—when your child figures out that he is responsible for his own education—and he is actually excited by that idea.

—PRHS Parent letter to the PRSD School Committee



We will regularly collaborate to balance content knowledge with Habits of Learning in a powerful learning environment. We will measure progress by locally created benchmark assessments and standardized tests

— PRSD Mission

CONCLUSIONS: KEY THEMES AND LESSONS IN STORIES FROM THE FIELD



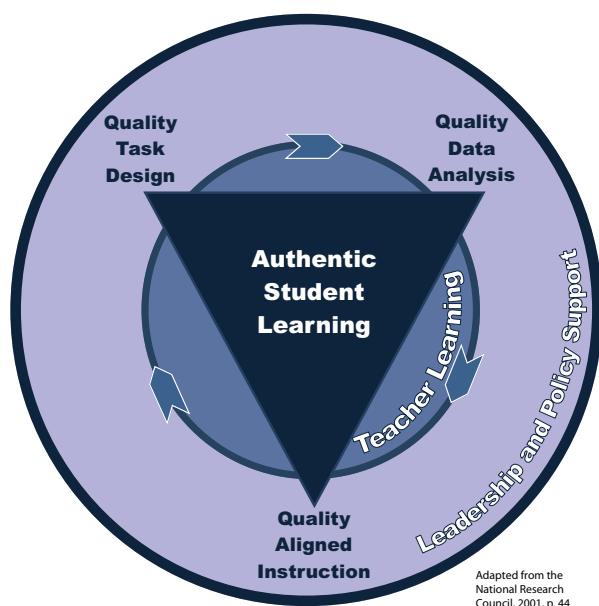
These three stories highlight different paths in developing high quality local assessment systems that incorporate performance assessments. While each story is characterized by distinct vision, actions, and outcomes, they share a common feature—each models how the QPA Essentials and the QPA Framework work together to improve student learning and teacher performance. School and district leaders seeking to create local assessment policies or to refine an existing policy can use these tools to guide decisions that reflect a high degree of assessment literacy. The following section briefly summarizes key policy lessons, which reiterate the QPA Essentials for performance assessment systems.

TECHNICAL QUALITY

Validity means an assessment measures what it was intended to measure. In the case of the Pentucket Regional School District, a high quality task design means assessments are aligned to essential local standards and are vertically aligned to help students master critical benchmarks. Beginning in elementary school, the Pentucket Regional School District's Habits of Learning enable teachers to create tasks that measure skills students need to succeed in college and the workforce, including the abilities to communicate effectively and think critically. Data from these assessments are used to drive instruction to ensure that all students have an opportunity to meet these standards and learn the skills they will need to be prepared for future endeavors.

Reliability means a group of teachers can agree on what a rating means and score it the same way. To ensure the integrity of the performance assessment process, schools and districts must take measures to ensure data is reliable and consistent. At Fenway High School, performance assessments measure students' ability to perform authentic tasks, including college-level writing in a Position Paper. The teachers at Fenway High School achieve reliability in the scoring of Position Papers through the use of common rubrics, anchor papers, and professional development time to analyze and score student work. A local system that is high in technical quality must provide training for teachers to develop consistent expectations and reliable scoring practices that do not vary significantly over time or across teachers, grade levels, or subject areas.

QPA Framework



PROFESSIONAL COMMUNITIES OF PRACTICE

All three stories emphasize the need for school and district leaders to provide time and resources for professional development that builds teachers' assessment literacy and skills. Teachers also need time to collaborate to ensure the quality and integrity of assessments. Student achievement will increase if schools and districts build an assessment-literate corps of teachers who are willing and able to transfer their learning to their colleagues in professional learning communities. At all three sites, teachers have led professional development sessions for colleagues in which they have presented assessment work and shared best practices. Similar models can be developed in schools and districts seeking to build internal capacity for assessment literacy.

School and district leaders have cultivated a high level of teacher participation and engagement by providing professional development time that is sustained over months, sometimes years. In the Pentucket Regional School District, for example, teachers and administrators spent 18 months developing the HOL rubrics before using them for instructional and assessment purposes. At Fenway High School, the Junior Review, Senior Institute, and Senior Position Paper have been refined in a community of practice for over 15 years. Time for Fenway High School's benchmark assessments is built into the professional development and student academic calendars to sustain continued improvement and effectiveness.

LEADERSHIP AND POLICY SUPPORT

The critical role of strong leadership and community support in performance assessment has been underscored in the work at all three schools. A clear vision for performance assessment was part of school and district policies. This vision has been at the heart of the design and implementation of the local assessment systems. The field stories demonstrate the various forms the investments of stakeholders and leaders may take—from political support built into the design of the school itself, to ongoing support from district and school leaders who provide ample time for professional development, to core groups of teachers who have a clear vision for and commitment to what powerful assessment can look like.

Pentucket Regional School District's and Fenway High School's experiences also reveal two important groups that can lend support to a local performance assessment system—students and their parents. When measures of learning are linked to important outcomes, students and parents can become advocates for the system because it measures competency and knowledge that is crucial beyond high school. Therefore, an important way to build local support for assessment policy is to link the assessments to meaningful, relevant measures beyond the classroom. Fenway High School's Senior Institute holds students to higher authentic standards of learning. The high level of buy-in, reflected by Fenway High School graduates who feel they have been adequately prepared for life beyond high school, is a testament to the importance of this work.

Key Lessons in Moving Forward with Quality Performance Assessment

In the 21st century, students will not be evaluated on their ability to complete multiple choice exams or memorize facts. Rather, they will be evaluated on their ability to think critically and creatively, communicate clearly, conduct authentic research using new media, and write persuasively and passionately. These skills call for the next generation of assessment—high quality performance assessments that measure how well a student transfers knowledge and applies complex skills. The need for these skills has been recognized at all levels, but it has yet to be realized. The untapped potential of performance assessments lies in the ability to demonstrate a higher level of competency and authentic learning. This potential will be realized when schools and districts strategically incorporate the QPA Essentials and focus on the technical quality attributes outlined in the QPA Framework.

As a country, we need to shift away from an education system in which students graduate based on “seat time” and course completion towards one based on graduation through demonstrating mastery and building content expertise in each discipline. If teachers collectively learn to design and implement performance assessments with technical quality, the policy door will be opened to teacher-designed performance assessments that count. Practitioner-developed assessment systems are the most effective way to ensure that students learn the skills outlined by the Common Core State Standards through curriculum alignment, authentic student learning, and meaningful assessments.

Some districts may be tempted to utilize commercially-developed performance assessments, thereby cutting short the time and resources needed for professional development. This decision would be a profound mistake. Teachers should participate in designing and using high quality performance assessments from the beginning. They should be given support and professional training to develop the assessment literacy needed to conduct assessments with technical quality. Without such central teacher involvement and professionalism, it is unlikely the next generation of assessments will promote the desired levels of student achievement and authentic student learning.

Schools and districts should appreciate both the short and long term benefits of a local performance assessment system. In the short run, a performance assessment system can provide better information about what students know and can do, provide timely data to shape instruction, increase student and teacher engagement, and help transition to the Common Core State Standards and the assessment system of tomorrow. In the long run, performance assessments can better prepare students for success as lifelong learners by shifting the focus to critical thinking, problem solving, and communication skills that will serve them well in college and throughout careers and civic life.

We hope to contribute to a conversation that leads to increased use of performance assessments and increased ownership of assessment practice at the local level, with the goal that students learn what matters most.

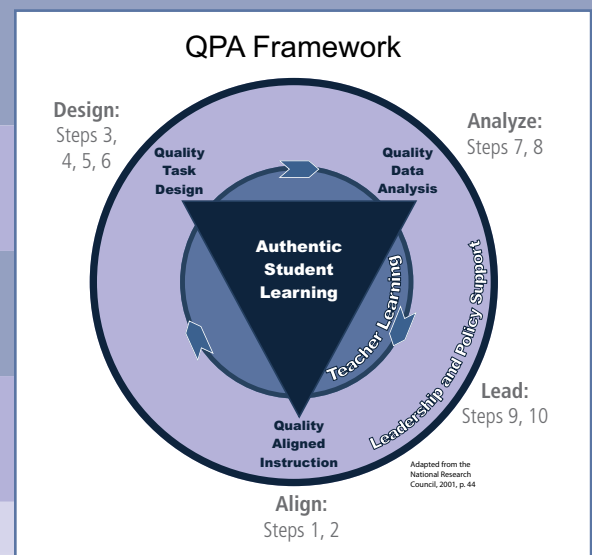


In the 21st century, students will not be evaluated on their ability to complete multiple choice exams or memorize facts. Rather, they will be evaluated on their ability to think critically and creatively, communicate clearly, conduct authentic research using new media, and write persuasively and passionately. These skills call for the next generation of assessment—high quality performance assessments that measure how well a student transfers knowledge and applies complex skills.

Steps to Quality Performance Assessment

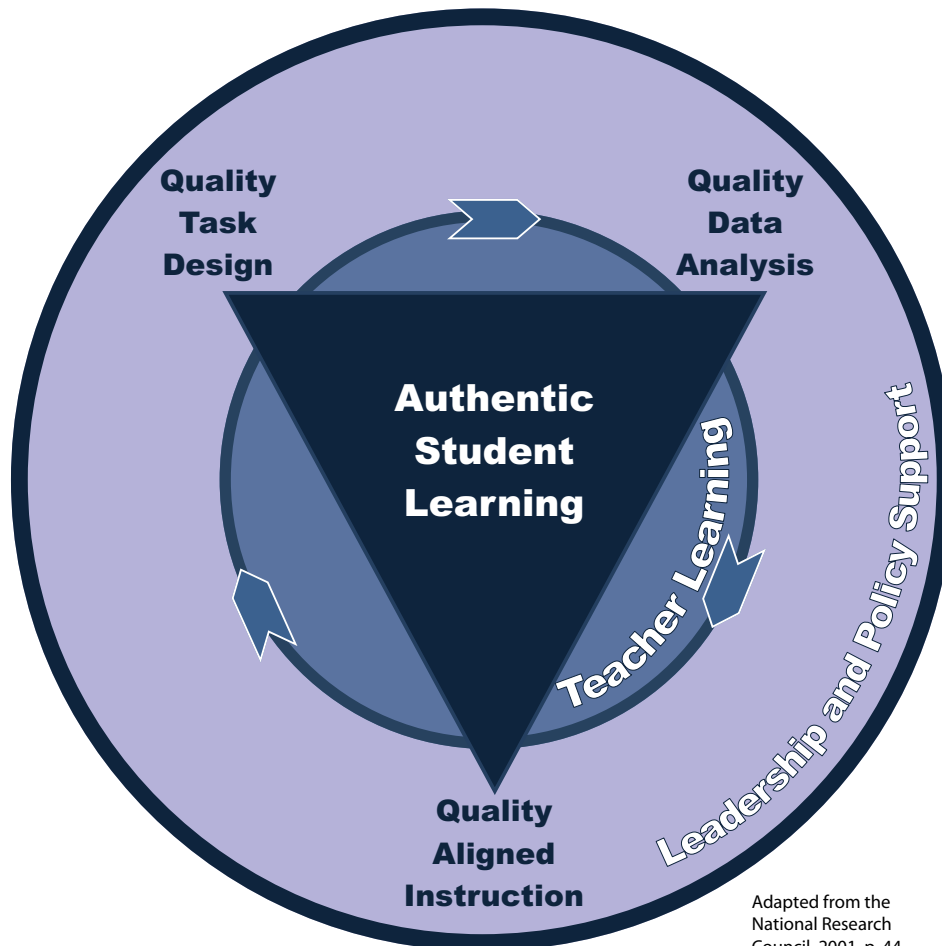
The purpose of the Steps to Quality Performance Assessment (QPA) is to illustrate the QPA Framework and provide guidance to teachers and administrators on the many possible entry points for engaging in this work.

1. Determine graduation and promotion requirements, essential learnings, and/or habits of mind and work that focus the school on the most important standards for students. Analyze course syllabi and assessments for alignment to prioritized standards.
2. Determine whether all students have learning opportunities and access to a rich and rigorous curriculum by aligning school structures and curriculum.
3. Ensure that the content and complexity of each assessment is appropriate for the assessment grade-level, based on the school's established content sequence and grade level standards.
4. Engage teachers in the design of performance tasks using clear criteria, agreed upon expectations, and processes that measure complex skills in multiple modalities. The tasks should be transferable to new situations and meaningful to students. These tasks should focus on authentic (real world) learning whenever possible; they should engage students, and provide opportunities for ownership and decision making in real world situations.
5. Provide a sample of student work at each performance level illustrating work at that level for each assessment.
6. Evaluate the use of universal design principles for each assessment (e.g., language clarity, use of white space and graphics) to ensure that all assessments are usable and effective, and that students have full access to the assessment.
7. Determine whether outcomes on assessments and the interpretations made about learning are closely related to student outcomes on other measures of the same standards.
8. Collect evidence to document consistency in scoring and calculate a reliability score for each assessment.
9. Document and adopt the local assessment policy through a process that builds the political will and support of all stakeholders.
10. Design professional development in communities of practice that supports all teachers in effectively implementing the policy.



APPENDIX: THE QUALITY PERFORMANCE ASSESSMENT (QPA) FRAMEWORK

QPA Framework



The purpose of the Quality Performance Assessment (QPA) Framework is to provide guidance to teachers and administrators on how to design a performance assessment system.

QPA has developed the following definition of performance assessment:

Quality Performance Assessments are 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.

QPA Framework Elements

The following elements guide the work of creating quality performance assessments:

TECHNICAL QUALITY

- **Aligned instruction**—To gain knowledge and skills, all students need instruction, based on college and career ready standards, that is accessible to their learning strengths and needs.
- **Task design**—Valid assessment tasks include determining appropriate levels of content and cognitive complexity, setting clear criteria for success, and ensuring accessibility for all students.
- **Data Analysis**—By examining student work and score data, practitioners interpret important information for planning future instruction and assessments.¹²

AUTHENTIC STUDENT LEARNING

Authentic student learning is the goal of this iterative cycle. QPA defines authentic learning as learning that is meaningful to students in which complex skills and content are embedded and transferable to new situations. Such learning can be assessed in multiple modes. Authentic learning engages students and provides opportunities for ownership and decision-making in real world situations. Because practitioner-developed performance assessments¹³ are created by those closest to the learner, they effectively guide and assess authentic learning.

TEACHER LEARNING IN COMMUNITIES OF PRACTICE

Teacher learning occurs when teachers engage with the three elements of quality assessment at the vertices of the triangle—aligning instruction, designing assessments, and analyzing performance data. To implement performance assessments effectively, teachers must collaborate with colleagues to understand how students best learn content and skills aligned to the standards, how to design assessments to elicit evidence of student competency, and how to interpret the student work. Engaging in professional dialogue about aligned instruction, task design and analysis of student work creates a synergy and ensures the level of quality required for authentic learning through performance assessment.

LEADERSHIP AND POLICY SUPPORT

Leadership at the school and district levels is essential to student and teacher learning. Leaders must create and document policies that support performance assessment such

as including authentic learning as a measure in promotion and graduation requirements. For these policies to be implemented with technical quality, leaders must devote professional development time to building teachers' capacity to align, design and analyze performance assessments. Furthermore, engaging families and the community in discussions of the value of practitioner-developed performance assessments builds political support for assessment policies that include authentic learning.

Evaluating the Technical Quality of a Local Performance Assessment System

The QPA Framework describes processes that may be implemented over time to design a local performance assessment system with technical quality. Many aspects of the QPA Framework may also be integrated into an existing local assessment system, or used as an entry point without a comprehensive assessment overhaul. For example, a first step in improving the technical quality of a performance assessment could be to collect and analyze student work samples at each performance level.

The more detailed descriptions of the QPA Framework elements guide practitioners in designing and evaluating a local assessment system. Each element includes a list of questions for evaluating the level of technical quality. If, upon review, the answer to all questions is “Yes,” the system likely has strong technical quality. If the answer to any question is “No,” QPA provides tools, professional development modules, and coaching to support schools in achieving technical quality.

QUALITY ALIGNED INSTRUCTION

Teaching and assessment practices are interwoven. Instruction and performance assessments are aligned and based on essential learnings¹⁴ and appropriate national, state, district, and school standards. In addition, all students need instruction based on college and career ready standards. Once students have the opportunity to master these standards, assessments allow them to demonstrate what they know and are able to do. Answers to the following questions are used to evaluate the system's level of technical quality.

¹² The criteria are aligned with the three vertices of the National Research Council Assessment Triangle: Cognition, Observation, and Interpretation (National Research Council, 2001).

¹³ The QPA framework focuses on performance assessment, although other types of assessments would also be included in a local assessment system.

¹⁴ Essential learnings are the critical skills, knowledge, and dispositions that are prioritized as being essential for every student to master (Dufour et al., 2006).

- a. Are promotion and graduation requirements aligned to essential learnings and appropriate agreed upon standards?
- b. Do standards include 21st century skills, such as collaboration and communication through multiple modalities?
- c. Are teaching and assessment practices for each course or classroom aligned to essential learnings and standards?
- d. Do all students have adequate time prior to the assessment to build upon prior learning, and to both practice and master the essential learnings and standards being assessed?
- e. Are students in different levels of the same course assessed with common performance assessments?

QUALITY TASK DESIGN

Effective assessment development begins with clarity about what students at each grade level should know and be able to do. A common understanding among faculty about content and cognitive complexity in the grades they teach and adjacent grades guides the design of prompts and scoring tools. Documentation of the assessment design and a validation process build awareness of expectations, allowing appropriate performance levels to be set at each grade level. Answers to the questions listed below are used to evaluate the quality of task design.

- a. Do the assessments measure complex skills in multiple modes, transfer to new situations, and provide opportunities for student ownership and decision making in real world situations?
- b. Are the content and cognitive complexity for each assessment aligned with established content and skills sequences and/or grade-level standards?
- c. Are there clear criteria and descriptions of performance at each level and aligned rubrics to assess mastery of the standard(s)?
- d. Are all assessment documents clear and understandable for all students? Do they address all aspects of universal design for learning¹⁵, e.g., clear student directions, use of white space and font size, unambiguous graphics? Are there appropriate and alternative response formats to allow for reasonable and fair accommodations?
- e. Are benchmark samples of student work that clearly define and illustrate work at each performance level available to all stakeholders?

QUALITY ASSESSMENT DATA ANALYSIS

Quality assessment data analysis involves examining student work and score data for *technical quality*. This means assessments must be valid, reliable and provide sufficient evidence of learning. *Valid* means the assessment measures what it was intended to measure (both content and intended level of rigor). *Reliable* means a group of teachers (or scorers) agree on what a rating means and can score it the same way. Reliability is essential because assessment data leads to high and low-stakes actions and decisions. To ensure that all students are demonstrating mastery, sub-group performance should be examined for bias in score results. *Sufficient* means students have been given a complete opportunity to demonstrate mastery resulting in enough evidence of learning being collected. Without technical quality there will be no guarantee that an assessment system has evaluated student learning fairly and completely. Conclusions from the data analysis provide information to practitioners for planning future instruction and assessment. Answers to the questions listed below are used to evaluate quality data analysis.

- a. Is there professional development for scorers that uses scoring guidelines and benchmark student work samples?
- b. Is there a process for collecting scoring data and auditing the scoring process to ensure scores are consistent across administrations and raters?
- c. Is there a systematic process for analyzing scoring data for student subgroups?
- d. Does the assessment provide the information about mastery of standards/content for which it was designed?
- e. Do teachers and other faculty use performance assessment data to inform curriculum planning, instruction and (re-)design of assessments?

¹⁵ Universal design for learning requires that assessments are accessible in language and format for the broadest possible range of students (Rose & Gravel, 2010).

REFERENCES

- Conley, D. T. (2007). *Redefining College Readiness* (Vol. 3). Eugene, OR: Educational Policy Improvement Center.
- Costa, A. (2008). Chapter 2: Describing the Habits of Mind. In A. L. Costa & B. Kallick (Eds.), *Learning and Leading with Habits of Mind*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Dufour, R., Dufour, R., Eaker, R., & Many, T. (2006). *Learning by Doing: A Handbook for Professional Learning Communities at Work*. Bloomington, IN: Solution Tree.
- Duncan, A. (Ed.). (2010). *Beyond the Bubble Tests: The Next Generation of Assessments—Secretary Arne Duncan’s Remarks to State Leaders at Achieve’s American Diploma Project Leadership Team Meeting*: US Department of Education.
- Finn, C. E., & Petrilli, M. J. (2010). *Now What? Imperatives and Options for “Common Core” Implementation and Governance*. Dayton, OH: Thomas B. Fordham Institute.
- Hess, K. (2009). *Local Assessment Toolkit: Assessment Task Validation Tools - Criteria for High Quality Performance Assessment*. Retrieved from www.nciea.org/publications.
- National Research Council (Ed.). (2001). *Knowing What Students Know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Pentucket Regional School District, & Pentucket Regional High School. (2010). *Habits of Learning Powerpoint Presentation*. West Newbury: Pentucket Regional School District.
- Rose, D. H., & Gravel, J. W. (2010). Universal Design for Learning. In B. McGaw, P. Peterson & E. Baker (Eds.), *International Encyclopedia of Education* (3rd ed.). Oxford: Elsevier.
- Solis, J. (2011, March 07). Pentucket program impresses state education chairman. *Newburyport News*. Retrieved from <http://www.newburyportnews.com/local/x831629152/Pentucket-program-impresses-state-education-chairman>
- Stefanakis, E. (1997). *The Senior Institute Experience: Students’ Stories of their Culminating Year at Fenway Middle College High School*. Boston: Center for Collaborative Education.
- Stiggins, R. J. (1999). Teams. *Journal of Staff Development*, 20(3), 17-21.
- Tung, R., & Stazesky, P. (2010). *Including Performance Assessments in Accountability Systems: A Review of Scale Up Efforts*. Quincy and Boston, MA: Nellie Mae Education Foundation and Center for Collaborative Education.
- Webb, N. L. (1997). *Criteria for Alignment of Expectations and Assessments in Mathematics and Science Education* (pp. 47). Madison, WI: National Institute for Science Education
- Wood, G. H., Darling-Hammond, L., Neill, M., & Roschewski, P. (2007). *Refocusing Accountability: Using Local Performance Assessments to Enhance Teaching and Learning for Higher Order Skills*. Washington, DC: Forum for Education and Democracy.

ACKNOWLEDGEMENTS

- **Report Authors:** Christina Brown and Pascale Mevs
- **QPA Team:** Laurie Gagnon, Rosann Tung, Susan Westlund
- **QPA Interns:** Maggie Ward, Erica Schwartz, Jennifer Hanson, Carolina Prieto, Stephanie Lloyd
- **CCE Editorial Group:** Robert Frank, Dan French, Leah Rugen, Alethea Pratt, Meg Anderson

We wish to thank the following people for their contributions to this paper and the QPA Initiative:

- Dedicated faculties and leadership of the schools featured in the Executive Summary and full paper—Cape Cod Lighthouse Charter School, Pentucket Regional Middle School and High School, Fenway High School
- Dedicated faculties and leadership of participating schools
- Dr. William Hart, Assistant Superintendent of the Pentucket Regional Schools
- Dr. Karin Hess, The Center for Assessment, QPA technical advisor and external reviewer
- State departments of education in Massachusetts and Rhode Island

The production of this white paper was generously funded by the Nellie Mae Education Foundation.

For accompanying tools see the QPA website: www.qualityperformanceassessment.org.

The mission of the Center for Collaborative Education (CCE) is to transform schools to ensure that all students succeed. We believe schools should prepare every student to achieve academically and to make a positive contribution to a democratic society. CCE partners with public schools and districts to create and sustain effective and equitable schools.

Citation: Brown, C. and Mevs, P. (2012). *Quality performance assessment: Harnessing the power of teacher and student learning*. Boston, MA: Center for Collaborative Education.

Copyright 2012 by the Center for Collaborative Education, Boston, MA, and Nellie Mae Education Foundation, Quincy, MA. All rights reserved.



Quality Performance Assessment: A Guide for Practitioners

This is a comprehensive guide aimed at assisting practitioners with the creation and implementation of quality performance assessments aligned to the Common Core State Standards. The QPA Guide builds on the work laid out in this paper. It is organized around the QPA Framework, providing tools and stories from the field for each Framework element. (In preparation for release in Summer 2012)

For more information about the QPA Guide, visit our website www.qualityperformanceassessment.org or contact us at QPA@ccebos.org.

Q

P

A

www.qualityperformanceassessment.org