nellie mae EDUCATION FOUNDATION

Center for Collaborative Education

Quality Performance Assessment

FEATURED STORY FROM THE FIELD

CAPE COD LIGHTHOUSE CHARTER SCHOOL: QUALITY DATA ANALYSIS AND COLLABORATION

Christina Brown and Pascale Mevs

FEBRUARY 2012

www.qualityperformanceassessment.org

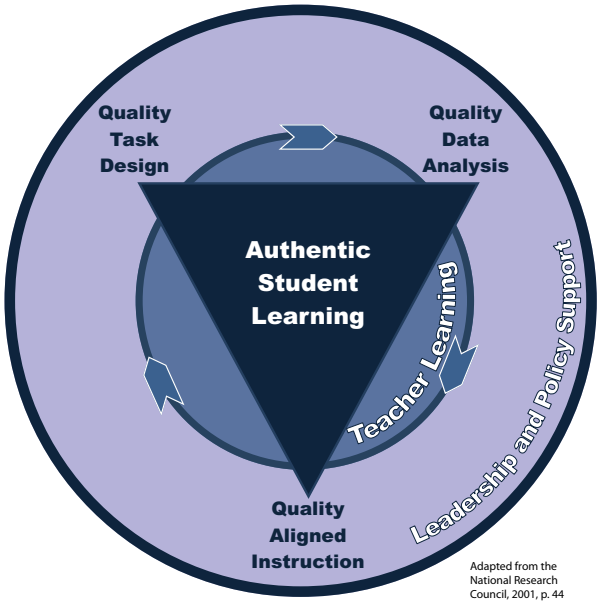
This story from the field is an excerpt from the report, *"Quality Performance Assessments- Harnessing the Power of Student and Teacher Learning."* The report is part of a project funded by the Nellie Mae Foundation to demonstrate the power of performance assessments to transform student learning and teacher practice. By sharing these stories, Quality Performance Assessments (QPA) aims to assist other schools in the process of adapting and implementing performance assessment systems. The research that produced these three stories from the field was conducted in participating QPA Network schools from 2009-2012. These stories outline key lessons learned from implementing performance assessments and formed the basis for the QPA Framework and the Steps to Quality Performance Assessment (Steps are located on last page of this document). QPA used qualitative research methods including: interviews, observation, and document review to detail each school's process of implementing performance assessments.

Each story takes an in-depth look at one of the following aspects of the QPA framework.

- Collaborate: Collaborating in professional learning communities of practice focused on performance assessment
- Align: Aligning assessments with local, state, and Common Core/national standards supporting college and career readiness
- Design: Designing performance assessments to allow students to demonstrate mastery
- Analyze: Analyzing data from performance assessment to drive instructional and programmatic decisions
- Lead: Leading the change process focused on policy, professional development, and community engagement

For more information about the work of QPA or to view the full report, please visit the QPA website at www.qualityperformanceassessment.org.

QPA Framework



Adapted from the National Research Council, 2001, p. 44

TABLE OF CONTENTS

Note: Page numbers do not start at 1 because this story is excerpted from the full paper, *Quality Performance Assessment: Harnessing the Power of Student and Teacher Learning,* to illustrate the QPA Framework elements: Analyze and Collaborate. This particular story outlines the work of educators at Cape Cod Lighthouse Charter School in Orleans, Ma, in analyzing student data and collaborating as a professional learning community as they implement the assessment validation process.

ACKNOWLEDGEMENTS

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 paper—Cape Cod Lighthouse Charter School, Pentucket Regional Middle School and High School, Fenway High School
- Dr. Karin Hess, The Center for Assessment, QPA technical advisor and external reviewer
- State departments of education in Massachusetts and Rhode Island

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

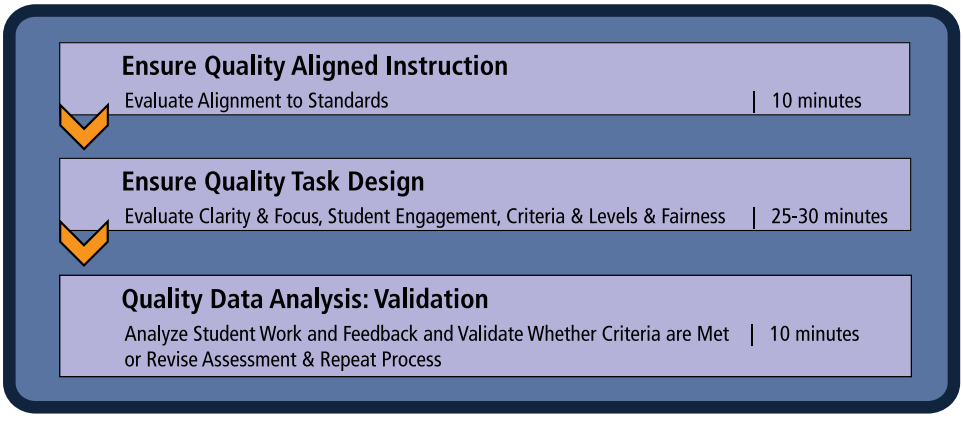
Next, the team spends 25-30 minutes evaluating the assessment's clarity and focus, the opportunities provided for student engagement, the appropriate use of rubrics or scoring guides, and the degree of fairness. The validation team looks for evidence the assessment is linked to instruction, requires 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.

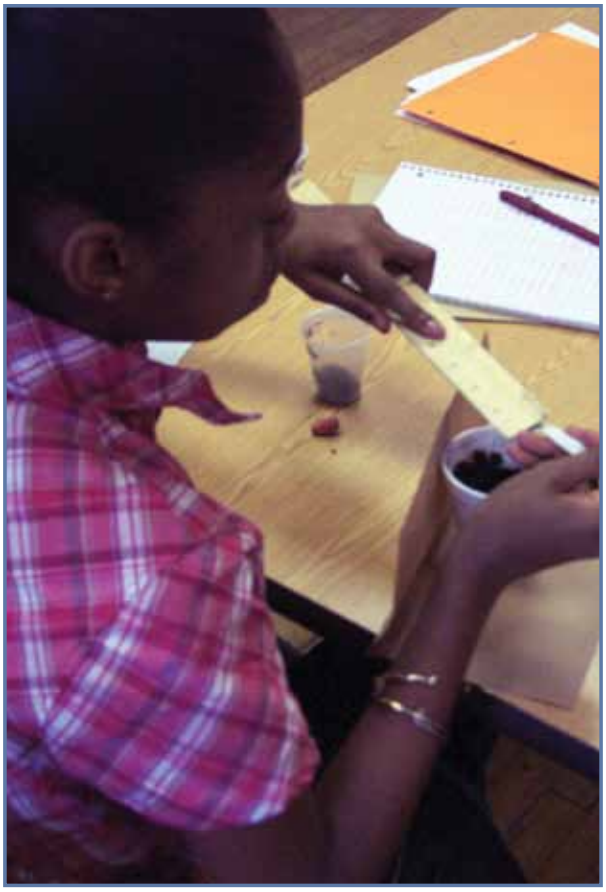
CCLCS's local assessment system includes a series of benchmark performance assessments students must pass in each grade level. Reviewing student scores on these benchmark assessments and comparing them to how students perform on other indicators—including grades, standardized tests, and diagnostic tests 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.

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.