

MOVING FROM GOOD TO GREAT IN NEW HAMPSHIRE:

Performance Assessment of Competency Education (PACE)

This overview summarizes the New Hampshire Department of Education's (NHDOE's) Performance Assessment Pilot, approved by the US Department of Education on March 1, 2015 (revised January 2016)



OVERVIEW

New Hampshire was awarded permission from the U.S. Department of Education in March 2015 to pilot an accountability system designed to support deeper learning for students and powerful organization change for schools and districts. The accountability pilot, referred to as **Performance Assessment of Competency Education or PACE**, is grounded in a competency-based educational approach designed to ensure that students have meaningful opportunities to achieve critical knowledge and skills. The recently passed **Every Student Succeeds Act (ESSA)** includes the "**Assessment and Accountability Demonstration Authority**" which was modeled after New Hampshire's PACE initiative and extends this opportunity for up to seven states. There are several key components to



Performance Assessment of Competency Education or PACE, is grounded in a competency-based educational approach designed to ensure that students have meaningful opportunities to achieve critical knowledge and skills.

this approach that NH DOE believes will help us achieve better results for all students:

- ✓ Explicit involvement of local educational leaders in designing and implementing the accountability system,
- ✓ Intense and reciprocal support on behalf of the NH DOE for local districts involved in this initiative that will include technical, policy, and practical guidance,
- ✓ Use of competency-based approaches to instruction, learning, and assessment which can best support the goal of college and career readiness for all students, and
- ✓ Use of instructionally-relevant, high-quality performance-based assessments, alongside periodic administration of Smarter Balanced and SAT assessments of state standards in math and English language arts (ELA), for the purpose of tracking and reporting the progress of students, schools, districts, and educators.

In a competency-based system, students' opportunities are judged by the outcomes they achieve and not by "inputs" such as seat time. Therefore, students must achieve identified learning targets before moving on to the next goals and/or graduating from high school. If they do not, school

districts are expected to work with families to support additional learning opportunities to ensure that students have legitimate opportunities to master the necessary knowledge and skills.

High quality performance assessments play a crucial role in the NH PACE system because of the need to measure the depth of student understanding of these complex learning targets. Performance assessments are used both to inform teachers and students of how the learning activities are working and what might need to be adjusted (formative) along with serving to help document what students have actually learned (summative).

LEARNING ALONG THE WAY

The DOE is aware of the well-known challenges with implementing performance assessments as part of accountability systems. A key challenge is building the local capacity necessary to raise the level of student learning, improve local performance assessments, and support local accountability determinations. One of the most important lessons we have learned in the pilot program to date is that the cross district collaboration provides a high level of professional learning opportunities. Having districts join networks of districts focused on similar goals has turned out to be an especially effective and sustainable strategy.

One of the DOE's major areas of focus is scaling such efforts to all NH schools. The current NH PACE accountability system is based on a voluntary proof of concept pilot with high-capacity schools. Improving chronically low-performing schools will be a priority over time. The DOE is committed to supporting the development of local leadership and capacity to help low performing schools implement the NH PACE system with fidelity.

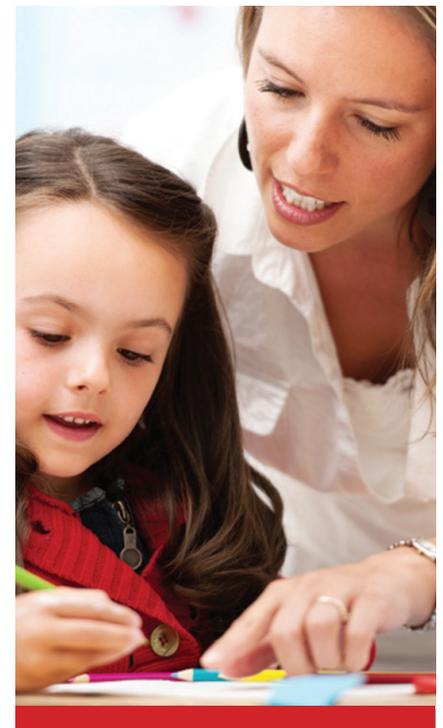


High quality performance assessments play a crucial role in the NH PACE system because of the need to measure the depth of student understanding of these complex learning targets.

One of the most important lessons we have learned in the pilot program to date is that the cross district collaboration provides a high level of professional learning opportunities.

KEY ASPECTS OF PACE

Who	<p>Participating School Districts for 2015-2016 include: Sanborn Regional School District, Rochester School District, Epping School District, Souhegan School District, Concord School District, Pittsfield School District, Seacoast Charter School, and Monroe School District. Several other school districts are currently building their capacity to become fully participating NH PACE districts in subsequent years.</p> <p>Importantly, the “who” includes essentially all students in each district. Instructional and assessment accommodations are available for students with disabilities as well as students for whom English is not their native language. A fundamental value of NH PACE is that the system should be designed to maximize the learning opportunities for each individual student.</p>
When	<p>NH PACE began as a two-year pilot. The Assessment and Accountability Demonstration Authority under the Every Student Succeeds Act would permit a three-year demonstration and a two-year extension contingent upon a successful three-year program.</p>
Pilot	<p>This federally approved pilot is not a “waiver” and does not otherwise change the requirement that all districts participate in the Smarter Balanced (and now SAT) assessment. Similar requirements will exist for other states that hope to be awarded “demonstration authority” status under the Every Student Succeeds Act.</p>
Learning	<p>NH PACE is a learning system designed to capitalize on the latest advances in understanding of how people learn.¹ The goal is to structure learning opportunities that allow students to grapple with gaining meaningful knowledge and skills at a depth of understanding that they can transfer to new real-world situations. As a coherent system, NH PACE is designed to foster positive organizational learning² and change by supporting the internally-driven motivation of educators instead of the all-too-common top-down accountability approaches where the goals and methods of the accountability system are defined at the state or federal levels and districts are simply expected to comply.</p>



NH PACE is a learning system designed to capitalize on the latest advances in understanding of how people learn.¹

¹ Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school (Expanded Edition)*. Washington, DC: National Academy of Sciences.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29, 7, 4-14.

² Elmore, R. F. (2004). Moving forward: Refining accountability systems. In Fuhrman, S. H. & Elmore, R. F. *Redesigning accountability systems for education* (pp.276-296). New York, NY: Teachers College Press.

Fullan, M. (2001). *Leading in a culture of change*. San Francisco: Jossey Bass.

Pink, D. H. (2009). *Drive: The surprising truth about what motivates us*. New York, NY: Riverhead Books.

Assessments The **core of the NH PACE assessment system is locally-developed, locally-administered performance assessments** tied to grade and course competencies determined by local school districts. Additionally, in each grade and subject without a state assessment (a total of 17 subjects and grades), there is **one, common complex performance task** administered by all participating districts. This common assessment is **NOT a state test!** Rather, it is developed collaboratively among the participating districts and is used to ensure that each teacher's evaluation of student performance is **comparable** to the evaluations made by other teachers. Finally, Smarter Balanced is administered in grade 3 (English language arts), 4 (math), and grade 8 for both ELA and math. The SAT is administered to all grade 11 students. **In other words, "state" assessments are administered in only 6 grades/subjects and local assessments in 17.**

"Over-testing" **NH PACE eliminates over-testing!** Teachers use every assessment except the six "state" assessments to inform day-to-day student learning, provide data for student grades, and help determine students' levels of competency. Legitimate concerns about over-testing arise when students take tests with no direct benefit to the students or the school but are required for accountability purposes. NH PACE avoids the problem of over-testing because the same performance assessments used for local grading also serve accountability purposes.

Accountability All parts of the assessment system described above count in state and federally required "accountability determinations." Smarter Balanced assessments and SAT are, therefore, administered less frequently in NH PACE districts.

"Guardrails" Districts must demonstrate readiness to participate in the pilot and must make certain commitments to continue with the pilot. **NH PACE districts must:**

- Have adopted the state model graduation and grade-span competencies and developed a coherent set of K-12 course and grade competencies.
- Have demonstrated the leadership and educator capacity to participate effectively in the pilot.
- Have developed or be close to completing the development of a comprehensive assessment system designed to appropriately measure student learning of required competencies.
- Be willing to participate in a peer and expert review process where they submit their system of performance-based assessments for evaluation based on clear and rigorous criteria.
- Administer Smarter Balanced or SAT in high school at least once at each grade span (e.g., 4, 8, and 11) to serve as both an internal and external audit regarding school and district performance.



The **core of the NH PACE assessment system is locally-developed, locally-administered performance assessments** tied to grade and course competencies determined by local school districts.

NH PACE avoids the problem of over-testing because the same performance assessments used for local grading also serve accountability purposes.

Reciprocal Compared with more compliance-oriented state accountability approaches, **NH PACE is a true partnership among the participating districts, the NHDOE, and technical support consultants.** The decisions regarding NH PACE are made collaboratively among the districts and the state, with each partner expected to live up to its responsibility.

Support Significant resources from regional and national foundations have been critical to the success of this pilot.

TECHNICAL QUALITY AND QUALITY ASSURANCE

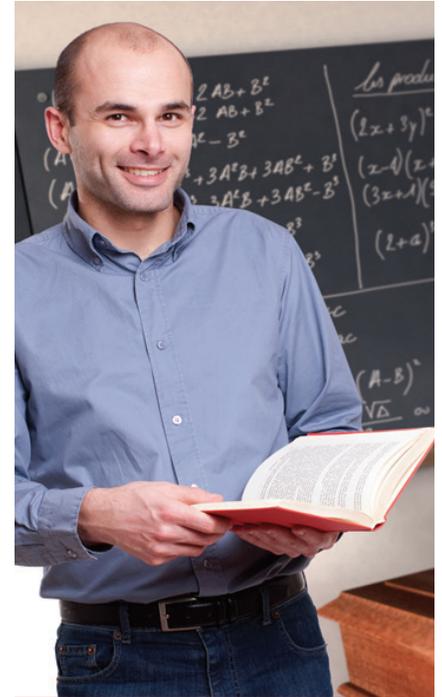
NH PACE relies on **local educators**, working collaboratively, to make important decisions about student performance. In addition to the quality of the specific performance assessments used in PACE, there are several other key technical details that are addressed as part of the NH PACE design, including:

1. Which assessments comprise the NH PACE system?
2. How is the quality of the assessments assured?
3. How does NH PACE support fairness and accessibility concerns?
4. How are NH PACE assessments scored and how is scoring quality ensured?
5. How is comparability evaluated and monitored?
6. How are “annual determinations” produced?

1. Which assessments comprise the NH PACE system?

Teachers give formative assessments many times a week. They may be homework assignments, quizzes or just questions asked during class. In competency-based schools, teachers use this formative information to determine when students are ready to demonstrate their mastery of various knowledge and skill areas (competencies).

NHDOE is not involved in those local assessments except to offer technical support when requested. However, NHDOE brings NH PACE districts together to create the high-quality common performance assessments used by all NH PACE districts. While these common assessments serve as an anchor for evaluating comparability across districts, teachers also use these assessments for documenting student competency and grades as well as benchmarks for local assessments. Finally, Smarter Balanced and SAT assessments are administered in select grades to help support comparability evaluations within PACE and between PACE and non-PACE districts.



NH PACE relies on **local educators**, working collaboratively, to make important decisions about student performance.

2. How is the quality of the assessments assured?

NH PACE employs five layers of quality control to ensure and continue to improve technical quality of the assessments used in the system.

- ✓ **Professional Learning and Collaboration:** Teams from each NH PACE school participate in high-quality workshops called Quality Performance Assessment Institutes. Then teams of teachers from all NH PACE districts collaborate and challenge one another in creating the performance tasks to be used as assessments. Teams for each content area in each grade follow a disciplined process of task development that includes “think aloud” protocols to try out the tasks before they are piloted and a small-scale field test to evaluate and refine task quality and to identify anchor papers to support reliable scoring.
- ✓ **Content Area Leads:** The NH DOE provides content experts in each of the three content areas (ELA, Math, and Science) to facilitate the process of developing the NH PACE common performance assessments and ensure that all NH PACE common tasks meet key design principles.
- ✓ **Center for Assessment Technical Review:** The National Center for the Improvement of Educational Assessment (Center for Assessment), a nationally-respected nonprofit assessment firm and our key technical partner, is responsible for conducting technical reviews of each NH PACE common task. The Center for Assessment bases its reviews on the High Quality Assessment Review Tool, a review protocol that has been used in multiple states and districts. The review tool distills key aspects of assessment quality from authoritative sources such as the Standards for Educational and Psychological Measurement,³ but presents the information in ways useful to local educators. The Center for Assessment provides feedback, particularly concerning the complexity of thinking required of students. Minor revisions are handled by the Content Area Leads, while more significant revisions must go back to the teacher teams and be resubmitted to the Center for Assessment.
- ✓ **New Hampshire DOE Review:** The project leaders from the NH DOE provide the final review and signoff on all NH PACE common assessment tasks.

NH PACE employs five layers of quality control to ensure and continue to improve technical quality of the assessments used in the system.

3. How does NH PACE support fairness and accessibility for all students?

Competency-based education is a purposeful approach for ensuring that all students, from the most advanced to the most challenged, move on only when they have mastered critical knowledge and skills at each level. We use the Universal Design for Learning (UDL)⁴ framework to minimize accessibility barriers and ensure that all students are able to show what they know and what they don't know.

Instruction and assessments based on UDL address the needs of the broadest range of students while avoiding

³ American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME) (2014). *Standards for Educational and Psychological Tests*. Washington, DC: AERA.

⁴ Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal Design Applied to Large Scale Assessments (Synthesis Report 44)*. Minneapolis, MI: University of Minnesota, National Center on Educational Outcomes. Retrieved September 5, 2015, from <http://www.cephd.umn.edu/NCEO/onlinepubs/synthesis44.html>.

potentially distracting information such as unnecessary wordiness and/or visually cluttered page layouts. Even with the use of UDL approaches, some students will need accommodations and related supports to be able to meaningfully participate in instruction and assessment.

While this is a civil rights issue, it is also important to the comparability goals of the PACE project that participating districts employ a common approach for using accommodations with the PACE common tasks. Therefore, participating districts and the leaders from NH DOE created PACE accommodations guidelines that define accommodations for students with disabilities and English language learners. This guide closely follows accommodations guidelines previously used for the NECAP assessments and now used for the Smarter Balanced assessments.

4. How are NH PACE assessments scored and how is scoring quality ensured?

Scoring of the NH PACE common assessment tasks follows a rigorous protocol developed collaboratively by our technical consultants and district leaders to ensure a high-degree of scoring consistency within each district and as a first step for ensuring cross-district comparability. This scoring framework identifies “anchor papers” (prototypical student work at each competency level) and monitors consistency to assure that teachers are scoring in a calibrated process.

Leaders and teachers in each district determine how to score their local competency-based assessments. The NH DOE PACE leaders do not layer additional requirements on local districts beyond what is necessary to document the quality of the results. However, many districts use the NH PACE common assessment approach as a model for their own local assessments as well.

5. How is comparability evaluated and monitored?

A core value of NH PACE is ensuring that all students from all participating school districts have opportunities to learn the required knowledge and skills. One of the ways to evaluate these opportunities is to require all students to participate in the same assessment, but that is not the only way. The “same” is not the only way to define comparability. For example, students applying for a competitive music program play different songs on different instruments, but judges still must determine who should be admitted to the program. We accept that judges are able to weigh the different types of evidence to make “comparable judgments.” We accept this because we have great trust in expert judges when they apply shared criteria.

In NH PACE true psychometric comparability (i.e., “interchangeability” of student scores) across districts administering different systems of assessment is not the goal. Rather, the focus is on evaluating and ensuring that students across NH PACE districts are held to comparable expectations. Student work on a NH PACE common task judged to be at Level 2 in one district should receive similar judgments of quality in every other NH PACE district. Further, NH DOE requires that students in NH PACE districts are held to the same or higher expectations as students from non-PACE

A core value of NH PACE is ensuring that all students from all participating school districts have opportunities to learn the required knowledge and skills.

districts. For example, students deemed “proficient” in a particular grade or content area should reasonably be expected to perform as proficient across types of assessments.

Led by the technical experts at the Center for Assessment employing comparability evaluation techniques used in Queensland, Australia and Great Britain⁵, NH PACE teachers and leaders participated in a summer 2015 comparability workshop where the rigor and consistency of local scoring was evaluated. This was a complex logistical activity in which multiple student work samples had to be scored by two teachers, neither of whom could be from the school that generated the paper. One of the most impressive findings from the comparability workshop was that after reviewing nearly 400 independent student papers, randomly assigned pairs of teachers were able to reach a consensus score on all but three papers. This is a reflection of the quality and commitment of the teachers and is a validation of the quality of the scoring rubrics, the foundation of scoring consistency.

The cross-district analyses revealed some differences in scoring across districts that varied by grade level and subject area, but essentially all differences fell within expected variation for the size of the sample being used and the natural measurement error inherent in all assessments.

6. How are “annual determinations” produced?

Annual determinations are the achievement classifications (e.g., proficient, basic) of students made at the end of the year typically. These annual determinations have typically been derived from single, end-of-year assessments such as NECAP where students are classified into one of four achievement levels.

Comparability in scoring performance assessment tasks is important but the ultimate goal is that “annual determinations” are comparable across school districts. In fact, a key requirement of the NH PACE initiative is to be able to produce “comparable annual determinations” for participating NH PACE districts and between PACE and non-PACE districts.

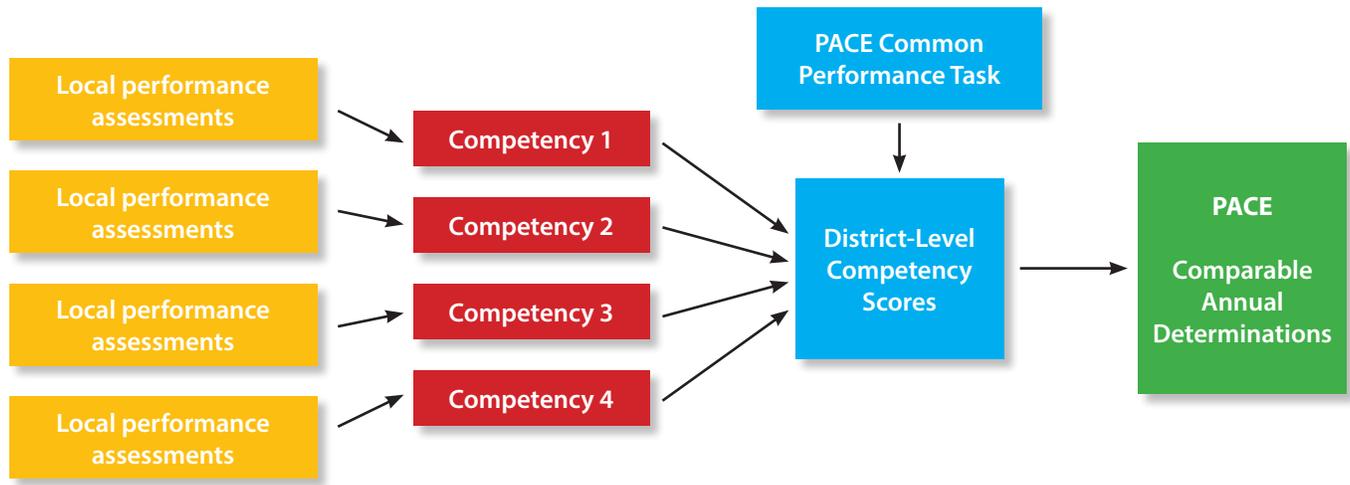
If a student in District X is called “Level 3” based on an examination of her work, then a student with similar levels of work produced in District Y should also be categorized as “Level 3.”

The NH DOE approach does this. NH PACE annual determinations are based on a combination of local summative assessments tied to specific competencies and the common PACE performance assessments as shown in the figure below. Our national technical advisory committee helped the NH PACE leadership team think through various considerations for designing a valid standard setting process.

The calibration of performance task scores discussed above is evidence that scoring among districts was comparable. The next step was to establish cutscores via a standard setting method on the distribution of competency scores collected in each district throughout the year.

⁵ Newton, P. E., Baird, J.-A., Goldstein, H., Patrick, H., & Tymms, P. (Eds.). (2007). *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.
Queensland Studies Authority. (2014). *School-based assessment: The Queensland assessment*. Queensland, Australia: The State of Queensland (Queensland Studies Authority).

Smarter Balanced/SAT in select grades



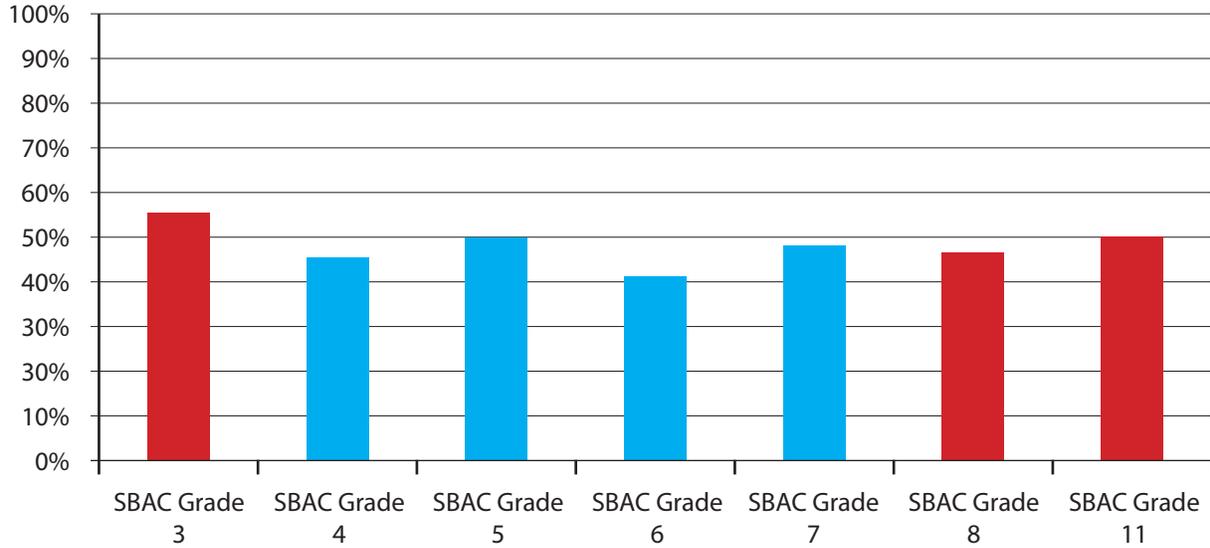
The purpose of the standard setting is to determine where in the score distributions the appropriate “cut points” lie for establishing achievement levels. To establish cut points we used an examinee-centered judgmental method called “contrasting groups”⁶. This standard setting method involves judgments from panelists about the qualifications of the examinees based on prior knowledge of the examinee. To implement this method for the NH PACE pilot, we asked NH PACE teachers to make judgments about which achievement level best described each of their students from the previous year. This process relied on the achievement level descriptors (ALDs) that were written by teachers on August 11, 2015. The ALDs for ELA and math were based closely on the ALDs used to guide the development and standard setting for Smarter Balanced. The science ALDs were drawn from the cross-cutting themes found in the Next Generation Science Standards (NGSS) and the NECAP science standards.

The subject and grade level specific ALDs were entered into an online survey where teachers could easily read the descriptions and match their former students to the appropriate achievement level. The power of the contrasting group (also known as teacher judgment) method is that it anchors teachers’ evaluations of student achievement in common descriptions of performance (the ALDs). The teacher directions for the contrasting group/teacher judgment survey made clear that teachers are to anchor their judgments in the actual ALDs and not some memories of students.

The contrasting groups standard setting methodology then compares the NH PACE scores with student placements into achievement levels in order to determine cut scores that would accurately classify the highest percentage of students into achievement levels. Since each district used a different constellation of local summative assessments, the analyses had to be conducted separately for each of the 17 grade/subject combinations for each of the four districts. The success of the cross-district calibration analyses provided the confidence that these results could be considered comparable. The overall results for ELA and math for the 2014-2015 are portrayed below. The annual determinations based on PACE assessments and Smarter Balanced were quite similar, indicating a high degree of comparability between PACE and non-PACE districts.

⁶ See: Cizek, G. J., & Bunch, M. B. (2007). Standard setting: A guide to establishing and evaluating performance standards on tests. SAGE Publications Ltd.

ELA: PACE Districts Percent Scoring at Level 3 & 4



Math: PACE Districts Percent Scoring at Level 3 & 4

