2017 No. 013

Formative Evaluation of New Hampshire's Performance Assessment of Competency Education (PACE)

Final Report

Prepared for:
New Hampshire Department of Education
101 Pleasant Street
Concord, NH 03301-3860

Prepared under:
Center for Innovation in Education
1648 McGrathiana Parkway, Suite 350
Lexington, KY 40511-0350
Contract Number UK-1652-16
(HumRRO No. S16-30)

Authors:
D. E. (Sunny) Becker
Arthur A. Thacker
Andrea Sinclair
Emily R. Dickinson
Anne Woods
Caroline R. H. Wiley

Date:
March 13, 2017

Formative Evaluation of New Hampshire's Performance Assessment of Competency Education (PACE)
Final Report

Table of Contents

Table of Contents (Continued)

Table of Contents (Continued)

List of Tables

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

List of Figures

Table of Contents (Continued)

Acknowledgments

We convey our deepest appreciation and sincerest gratitude to the educators, administrators, support staff, parents, and students in the PACE Tier 1 districts who participated in focus groups, interviews, and surveys that informed this evaluation: Concord School District (SAU 8), Epping School District School Administrative Unit (SAU) 14, Monroe School District (SAU 77), Pittsfield School District (SAU 51), Rochester School District (SAU 54), Sanborn Regional District (SAU 17), Souhegan School District (SAU 39), Seacoast Charter School (SAU 46), and White Mountains (SAU 35).

Formative Evaluation of New Hampshire's Performance Assessment of Competency Education (PACE) Final Report

Executive Summary

New Hampshire's Performance Assessment of Competency Education (PACE) is an assessment and accountability strategy designed to reduce the amount of, and reliance on, standardized testing by supplanting much of the traditional end-of-year summative testing with teacher developed performance assessment tasks. PACE was created to support deeper learning through competency education, and to be more integrated into students' day-to-day work than current standardized tests. The PACE pilot program represents a fundamental qualitative shift in the way accountability assessments are developed, administered, and used to promote teaching and learning.

In spring 2015, the U.S. Department of Education granted New Hampshire (NH) a waiver from specific requirements of the No Child Left Behind Act and then the requirements of the Every Student Succeeds Act (ESSA) as part of a demonstration pilot program.[1] Participating NH districts administer Smarter Balanced assessments in grade 3 English Language Arts (ELA), grade 4 Mathematics, and grade 8 ELA and math, as well as the SAT to all grade 11 students. In addition to local performance assessment tasks (hereafter local tasks), a common performance assessment task (hereafter common task) is administered in each grade and subject (ELA, math, and science) without a state assessment.

The Human Resources Research Organization (HumRRO) was awarded a contract to conduct a formative evaluation of the PACE system. This report is the fourth and final report in a formative evaluation of the PACE system in the Tier 1 districts, conducted from April 2016 to February 2017. The primary goal for this evaluation was to ensure that the PACE Leadership team has useful information to make decisions that advance the program's goals. As such, each (roughly) quarterly report built upon the previous report, capturing the state of the PACE system at the time, with cumulative descriptions of what was working well and specific feedback intended to help PACE Leadership make continuous process improvements.

HumRRO's June 2016 report (Becker, Thacker & Sinclair, 2016a) included observations of task development sessions and interviews with the eight PACE District Leads. The September 2016 report (Becker et al., 2016b) included interviews with PACE district leads, site visits to three districts and schools, and observations of various PACE events such as training of content experts, the PACE Tier 1 Summer Institute, the PACE Design Studio for Tier 2 and 3 Schools, and New Hampshire Department of Education's (NH DOE) Annual Educator Summit. The December 2016 report (Becker et al., 2016c) included observation of a Task Development Meeting, an interview with the PACE district Lead of the new Tier 1 district, a site visit to one district and its schools, attendance at monthly PACE Leads meetings, and a review of PACE standard setting and scoring. This final report includes detailed descriptions of activities conducted between late November 2016 and January 2017: visits to six districts and a sample of their schools, observation of a Task Development meeting, observation of a PACE Leads meeting, review of PACE standard setting and scoring data, and results from a teacher survey

[1] ED granted NH DOE a waiver extension on October 6, 2016.

administered to all teachers in Tier 1 districts. This report then consolidates this information, along with evidence cited in previous reports, into evidence for each of the nine success criteria.

A Brief Introduction to PACE

The PACE system relies upon locally developed, locally administered performance assessment tasks aligned with local district grade and course competencies. These local competencies and local performance assessments are aligned to the State Model Competencies, which, in turn, are aligned with national standards in each content area.

New Hampshire school districts must apply and demonstrate readiness and commitment before being allowed to participate in the PACE system. Districts enter via a three-tiered system, based on how fully they meet the requirements to implement PACE. Tier 1 districts have fully implemented PACE. Tier 2 districts implement competency-based education, but have not fully implemented PACE. Tier 3 districts are at a beginning stage. There are currently nine Tier 1 districts and they were the focus of the evaluation. Four districts joined PACE in 2014–15: Epping School District School Administrative Unit (SAU) 14, Rochester School District (SAU 54), Sanborn Regional District (SAU 17), and Souhegan School District (SAU 39). A second wave of districts became PACE Tier 1 districts in 2015–16: Concord School District (SAU 8), Monroe School District (SAU 77), Pittsfield School District (SAU 51), and Seacoast Charter School (SAU 46). In addition, White Mountains (SAU 35) joined as a Tier 1 district in the 2016–17 school year. SAU 35 was included in limited fall/winter 2016 evaluation activities. Because PACE replaces the Smarter Balanced assessments for several grade/subjects, the requirements for participation are rigorous. Districts must commit to administering a common task for every assessed grade/subject each year, plus they must agree to administer several local tasks. Students' scores on these tasks contribute to local student competency scores and feed into annual determinations. The tasks can often take several class periods to administer and a sample of papers must be double scored. Ensuring that the quality of all assessment stages, including developing, field testing, revising, administering, and scoring the performance tasks is sufficiently high requires a great deal of teacher professional development and a large time commitment for all participants.

Each common task undergoes a one-year pilot testing phase, with evidence-based revisions made after each round of pilot testing (the number of rounds determined based on the performance of the task), followed by an operational year. Administration of a pilot common task may occur in a subset of districts, but during the operational year, all Tier 1 districts administer the common task at the specified grade level. The common tasks must be administered in a standardized manner during the operational year to achieve comparability. After the pilot and operational years, these common tasks are available in a growing bank of tasks from which teachers can select to use as local tasks. Teachers may make modifications to the tasks at this time, including administering the task at a different grade level (Changing grade level would be primarily done in middle school science, where the curriculum is not consistent by grade across districts.). These tasks are also made available to Tier 2 and 3 to be used as teaching tools, once they are in the task bank after the operational year.

The PACE common tasks and local tasks are intended to be closely linked to classroom instruction. All the tasks, local and common, are teacher-designed to assess the specific competency targeted by lessons within the curriculum or unit of instruction. The tasks are not administered in a specific testing window, but instead come at the time during the year when it is most appropriate in the curriculum. Teachers know the content of the tasks well before administering them and the tasks are designed to test students' competency regarding

specifically taught content topics. There is no guessing what the tasks will cover in a given year. PACE tasks are complex and require deep understanding of the content. There are no multiple choice-questions on PACE tasks. Students write and revise, perform real-world applications of mathematics, or conduct science experiments to demonstrate their competencies. And, while PACE likely requires more testing time than Smarter Balanced, because it is so integrated into the curriculum, students often do not realize they are taking a test. Instead, they consider the PACE tasks to be another part of their daily classwork[2].

Framing the Evaluation

HumRRO was tasked with three evaluation goals:

- Evaluation Goal 1: **Refine and validate the PACE Accountability program's theory of change/theory of action**
- Evaluation Goal 2: **Provide formative feedback loops on key success criteria**
- Evaluation Goal 3: **Capture the "Story" of PACE**

Goal 2 included nine success criteria:

- Success Criterion 1: **Gaining clear commitment from local leadership**
- Success Criterion 2: **Building cross-district leadership and cross-district collaboration**
- Success Criterion 3: **Developing high-quality performance assessments**
- Success Criterion 4: **Successfully implementing common performance assessments**
- Success Criterion 5: **Providing training and calibration**
- Success Criterion 6: **Reaching successful rates of inter-rater agreement**
- Success Criterion 7: **Producing "comparable" annual determinations**
- Success Criterion 8: **"No harm" on the Smarter Balanced Assessments**
- Success Criterion 9: **Ensuring equity**

HumRRO provided interim reports, as well as informal feedback, organized around the goals and these success criteria to quickly provide ongoing feedback to the New Hampshire leadership during the course of the evaluation. The goals and criteria also served as major areas of inquiry for the final evaluation report.

At the onset of the evaluation, the theory of action/change was captured by three bullet points. They included:

- **"If we believe that all students must be college- and career-ready . . .**

[2] For a complete overview of PACE, see http://www.education.nh.gov/assessment-systems/pace.htm.

- **then, our system must advance students as they demonstrate mastery of knowledge, skills, and work study practices, . . .**

- **which requires a comprehensive system of educator and school supports."**

The bullet points are compelling, but do not lead directly to claims that can be investigated in a traditional validity argument. Our first task was to capture the goals of PACE and to map the success criteria onto a framework that could be used to organize and structure evidence collected regarding PACE's quality and validity. That framework is presented in Figure 1 as the theory of action/change for PACE.

Figure 1 includes four interim goals and a set of underlying claims that must be substantiated to attain each interim goal. Lack of support for any one interim goal may undermine subsequent goals. For example, if the tasks are not administered as intended (i.e., Interim Goal 3), then the validity of the scores is called into question (i.e., Interim Goal 4), regardless of how high inter-rater agreement and inter-rater reliability are among the scorers. While the interim goals are not entirely linear and dependent on each other (as they might be in a stricter interpretive argument for validation of an interpretation of assessment scores), this framework illustrates potential threats to the intended outcomes of the program. It also provides a common way of understanding how any potential threat within one of the interim goals might interact with others. The final evaluation report describes the various data collection activities and summarizes the evidence for each goal and its underlying claims and assumptions, thereby creating a validity argument for the PACE pilot program. In addition, the final report summarizes the successes PACE has achieved at this stage of implementation, concerns or issues that should be addressed, and conclusions and recommendations. Data collection was designed to include both qualitative and quantitative information from multiple stakeholders to triangulate and bolster the accuracy of the findings. Data collection methods included

- observations at major PACE events (e.g., task development and scoring sessions);

- classroom observations;

- interviews with students, parents, teachers, principals, and district leaders;

- surveys of teachers; and

- analyses of score data.

Collecting data from multiple stakeholders using multiple methods bolstered the accuracy of the inferences about PACE. It allowed us to capture the perceptions of the majority of PACE participants, and it allowed us to hear important minority opinions. Perhaps, most importantly, it allowed us to differentiate between the two.

NH Pace Chart: rev: 02.01.17

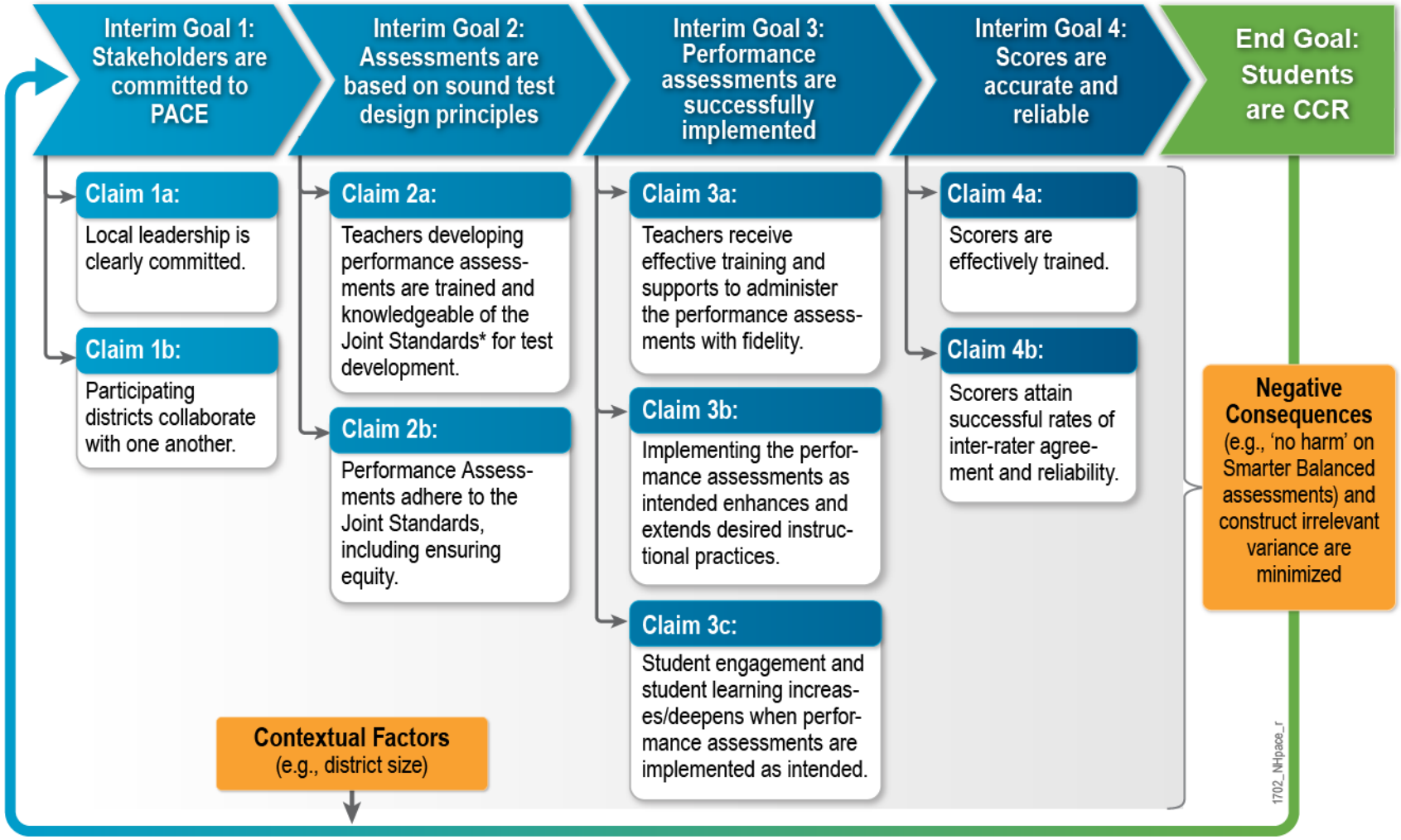


Figure 1. PACE theory of action/change.

** We understand that the PACE stakeholders are not test design experts and, therefore, that the AERA, APA, & NCME Standards are not firsthand knowledge for this audience. Consequently, our discussion with these stakeholders referred more generally to "high-quality assessment."*

Evaluation Activities

HumRRO conducted several data collection activities over the course of the evaluation, from April 2016 through February 2017. These included interviews with nine PACE District Leads; visits to schools in eight PACE districts to conduct interviews or focus groups with administrators, teachers, parents, and students, as well as classroom observations; observation of cross-district meetings including task development sessions and scoring and calibration sessions; participation in monthly PACE Leads Meetings; and review and analysis of scoring and calibration data. In addition, we administered a teacher survey to all teachers in Tier 1 districts, in part to help determine the generalizability of our findings from the teacher focus groups.

Evaluation

Interim Goal 1. Stakeholders are committed to PACE

We found strong evidence supporting Interim Goal 1. PACE participants overwhelmingly indicated that local leadership was highly supportive of the PACE initiative. There are several methods by which districts collaborate with one another, and participants report that collaboration is a major benefit of PACE membership. New collaboration mechanisms have recently been put in place to account for PACE expansion, but these have not yet been evaluated for effectiveness.

Claim 1a. Local leadership is clearly committed

The first testable claim from the Theory of Action is "local leadership is clearly committed" to PACE. For this claim, we gathered data from PACE District Leads, school administrators, and teachers. We include teachers because they are truly the most influential local leaders in the program. They develop the tasks and decide what is to be assessed. They also take the tasks back to their schools where they influence other staff members. The overwhelming majority of PACE participants reported high levels of commitment.

One of the most challenging requirements for the success of any educational intervention is securing buy-in from the major participants and leadership of classrooms, schools, and districts. PACE addresses this challenge in several ways. First, educators are in charge of nearly all aspects of the program. Teachers decide what is assessed, how it is assessed, and they even design the scoring rubrics. By placing the responsibility for creating the tasks on the primary users of the assessment data, PACE gives teachers more say in how their students will be assessed than in more traditional testing systems. Educators at all levels described ownership of the system as a major contributor to buy-in.

The second way PACE gains buy-in is by emphasizing the integrated nature of the assessments. Unlike end-of-year comprehensive statewide assessments, which sample from the past year's curriculum, PACE is targeted to the learning that is occurring at the time of administration. Since there is no specific testing window for PACE, and since the tasks are targeted to one broad curricular topic, teachers can administer the tasks when it makes the most sense. There is no need for intensive review during the weeks leading up to the testing window and no post-test slump between the end of the testing window and the end of the school year.

A third reason PACE participants are committed is that PACE replaces the Smarter Balanced assessments in the grade/subjects for which it is administered. As such, PACE provides an

alternative to an assessment that many New Hampshire educators regard as an interruption of their instruction that provides little useful information. PACE tasks require deep knowledge on the part of students. There is no chance of getting an answer correct by guessing. Students actually perform the tasks on which they are assessed, rather than answer questions about those tasks. PACE proponents describe the tasks as authentic and important. They often describe the benefits of PACE in terms of better preparing students for life beyond school. It is relatively easy to buy in to a program if you believe its methods and outcomes are better than what came before.

Claim 1b. Participating districts collaborate with one another

The second testable claim from the Theory of Action is "participating districts collaborate with one another." This claim is also supported in a number of ways. First, educators from all Tier 1 districts meet regularly throughout the year. They participate in task development sessions, professional development, scoring sessions, standard-setting, and other meetings. These cross-district meetings require that personnel from different schools work together to accomplish common goals. The meeting participants then implement the things they learn in their classrooms and share what they have learned with other educators within their school/district. A theme that emerged across the data collection activities is that teachers value and enjoy the opportunity for cross-district collaboration. They often refer to it as beneficial for their professional growth. They also describe it as useful for developing high quality common tasks and for calibrating the scoring of student work.

In-person meetings are just the beginning. The second way educators across districts interact is through the "LibGuide" system. This system is a repository for "all things PACE." It is a web-based repository for PACE tasks, rubrics, and shared resources. Teachers who implement common tasks early share their lessons and provide tips for smoother implementation among their colleagues. The teachers share book lists that are suitable for use in English language arts tasks. They share equipment lists for science labs, including locally available inexpensive options for commonly needed equipment. They also share guidance on the administration of the common tasks. Some commonly used documents include a guide for educational scaffolding, student-friendly rubrics, and principles of scoring student work.

Collaboration across districts is also accomplished by emailing the PACE coordinators and leadership. PACE teachers ask direct questions, some of which are answered individually, and some of which become group emails to eliminate potential common misunderstandings or misconceptions. If questions become common or concerning, they are addressed during in-person meetings and with guidance on the LibGuides.

Prior to the start of the evaluation, each district identified a PACE Lead to coordinate activities in the district and to communicate with PACE Leadership. Participation in monthly PACE Leads meetings is one venue for collaboration.

Over the course of the evaluation period, PACE implemented four new collaboration measures. The first was to name an overall curriculum coordinator to assist with PACE task development activities. This step was taken to (a) improve communication, (b) ensure common understanding of goals, instructions, and deadlines, (c) provide a master schedule earlier in the year, and (d) provide an additional resource for PACE participants.

Another new collaboration mechanism was the naming of multiple Content Leads (about 30 total) for each grade level and content area combination. These teachers were identified as leaders in PACE and were recommended by peers and ultimately selected by the PACE District

Leads to help coordinate subject/grade-specific activities. Most have been PACE participants and task developers since the beginning of the PACE pilot program. The Content Leads program allows PACE to build deep expertise among local educators without requiring all educators to attend every meeting and activity. The Content Leads helped PACE address the expansion of the program. They act as liaisons to the educators in their districts and also in a "buddy district," which might not have a Content Lead. This allows PACE to field smaller groups of teachers when a very large group would be unwieldy, such as during task revision workshops. Wordsmithing a common task can benefit from multiple voices, but there is a point of diminishing returns when too many group members provide input. The Content Leads help keep these kinds of in-person interactions small. This approach has the added benefit of reducing time that some teachers must spend outside the classroom in collaborative activities. The Content Leads take the information from workshops and other activities back to the districts. Educators who are not Content Leads can still provide information, including suggested revisions to common tasks and rubrics, via the LibGuide. In the districts without Content Leads, Teacher Representatives were identified to coordinate among local teachers.

The third new innovation is the "buddy district." Districts are now paired with other districts to promote collaboration. Districts with Content Leads are often paired with districts that do not have them. Newer PACE districts are typically paired with experienced districts. The Content Lead provides an opportunity for all participants to contribute to all aspects of PACE, in addition to the local tasks that all teachers develop. Buddy districts, as well as the other new collaboration initiatives, help PACE cope with expansion. As the program expands, these efforts become increasingly necessary to maintain the requisite levels of participation and ownership among PACE educators.

Finally, as of February 2017, PACE Leadership began inviting Tier 2 districts to attend monthly PACE Leads meetings. Observing these meetings will afford an opportunity to become more familiar with the PACE Tier 1 experience.

Interim Goal 2. Assessments are based on sound test design principles

We found strong evidence to support Interim Goal 2. The task developers (teachers) are well trained and thoughtful in the development of the tasks and scoring rubrics. They adhere to the central themes and major principles of the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014), even if they do not specifically reference them.

Claim 2a. Teachers developing performance tasks are trained and knowledgeable of the Joint Standards[3] for test development

In order for assessments to achieve their goals, providing valid and reliable inferences from students' scores, they must adhere to sound design principles. To build a sound assessment, we must understand the inferences that will be made from the scores, or put more simply, the ways we would like to use the assessment results. PACE scores serve three main purposes: (a) they provide progress checks on student learning at various points in the school year; (b) they inform teachers about students' competency related to specific knowledge, skills, and abilities; and (c) they are aggregated to provide an indication of school/district performance. Teachers routinely design assessments to check progress on the content they teach, and they

[3] We understand that the PACE stakeholders are not test design experts and, therefore, that the AERA, APA, & NCME Standards are not firsthand knowledge for this audience. Consequently, our discussion with these stakeholders referred more generally to "high-quality assessment."

did so prior to the PACE program. PACE adds the competency aspect, though many schools had implemented some form of competency education previously, placing the focus of the assessment on competency rather than progress or performance relative to peers. PACE is also used for accountability, as aggregate scores on PACE tasks replace Smarter Balanced scores for gauging school and district level performance. These new uses for the assessments require that the developers, in this case teachers, think through aspects of testing that they have not historically needed to consider. To do so, they must be effectively trained.

PACE teachers demonstrated high levels of assessment literacy during training sessions, scoring, and standards setting meetings. A large number of teachers who were trained and began the process of developing the common tasks are now highly knowledgeable about task design and provide a strong core of leadership for the program. While they do not typically reference the Joint Standards, they routinely discuss complex aspects of assessment design. They recognize and struggle with the dilemma between standardization and instructional and curricular freedom. They routinely discuss scaffolding as a method to give students access to assessment content, especially scaffolding for SWD. They recognize that scaffolding can represent both a benefit and a potential danger to the validity of student scores. They strive to ensure that the rubrics they design are well documented and can be consistently and accurately scored. When they design rubrics, they are careful to include only components for which students will provide appropriate evidence during the performance tasks. They routinely discuss the pros and cons of aspects of test design. For example, a common discussion point is whether to expand the content of a draft task. Broad tasks can address more content, but more discrete tasks may be better indicators of specific competencies.

The training model employed by PACE has allowed teachers to learn by doing, albeit with some assistance provided by assessment experts and personnel from the state department of education. They gain assessment literacy by encountering and dealing with assessment issues as they design, pilot, revise, administer, and score the tasks. When they have questions, they have expert help available, but they hold most of the decision-making power. This learn-doing model has been very effective for the teachers who started with PACE, but it takes a lot of time now to bring new participants up to speed. The experienced teachers have obtained considerable important knowledge, and orienting newcomers can be challenging. This will be an ongoing challenge as PACE expands to more districts.

Entailed in the theory of action for PACE is that teachers apply what they learn from developing high quality *common* tasks to the development of high quality *local* tasks. A survey of PACE teachers reveals the majority of teachers report that they have been able to apply what they have learned from their experiences developing *common* tasks to developing higher quality *local* performance tasks. This lends further support for claim 2a.

Claim 2b. Performance assessments must adhere to the Joint Standards, including ensuring equity

PACE teachers do not routinely reference the Joint Standards. They focus on solid assessment design and the tools they have been given. The common tasks used across all districts are highly scrutinized for potential biases or sensitive content. They undergo extensive editing and revision before and after they are field tested. Teachers follow detailed guidelines for administration, including guidelines for ensuring standardization and for promoting accessibility. They follow guidelines for providing accommodations. Students' work is double scored and scorers' consistency is verified. The common task is centrally rescored and used to adjust for any systematic scoring differences on local tasks by district. Well-established standard setting methods are used to classify students by performance categories. PACE results are compared

with an external reference assessment (Smarter Balanced). These efforts largely parallel the processes of large-scale testing companies that adhere to the Joint Standards and they contribute to a high quality assessment system.

The local tasks do not undergo the same levels of scrutiny as the common tasks. There were substantial differences in the quality and depth of the local tasks, by district, discovered during standard setting. According to PACE teachers, there were several factors that might account for the differences. Teachers learned late in the 2015-16 school year[4] that they were expected to keep documentation for a sample of students of local tasks conducted throughout the year. This may have caused some teachers to provide less than optimal student work samples. Teachers also have potentially differing levels of experience and expertise with competency based education, depending on when their district joined and how involved they were in PACE.

Improvements in communication and efforts to extend participation are expected to result in improved and more consistent local tasks. PACE teachers have indicated that they believe the common tasks and the local tasks are authentic measures of their students' achievement. Ensuring equity for all students is a challenge for any assessment system. When asked during interviews and focus groups whether PACE tasks were more or less accessible compared to Smarter Balanced, most teachers indicated that PACE tasks are more accessible. They described the embedded nature of PACE and the availability of the same accommodations students routinely received for classroom work as justification. Because PACE is embedded in instruction, students often do not realize that the PACE tasks are different from their regular class work. This helps make the tasks more accessible than an "assessment event." A few teachers, however, expressed concern that they might inadvertently impact the standardization, and consequently the measured construct, of the tasks (both common and local tasks) by providing too much accommodation or too much scaffolding. This is a common challenge for standardized tests and the concern led to the creation of an accommodations guide and a scaffolding guide to help teachers make informed and sound decisions about accommodations and scaffolding. Some teachers also expressed concern on the PACE teachers survey about the accessibility of PACE tasks. Contextual information gathered from teachers' open-ended comments on the survey, and also from focus groups, indicates that some teachers believe the reading and writing demands of PACE tasks are quite high and could limit accessibility for some students.

Interim Goal 3. Performance Assessments Are Successfully Implemented

We found strong evidence to support Interim Goal 3 for the common tasks. For the system overall, we found considerable evidence that the training and support are adequate and that PACE has had a substantial positive impact on both teaching practice and student learning. There was insufficient evidence to fully evaluate the quality and implementation of the local tasks at this time.

Claim 3a. Teachers receive effective training and supports to administer the performance assessments with fidelity

Most teachers report that their training is adequate for administering the PACE tasks. Most teachers report that their school's administration provides them with the resources and supports they need to effectively implement the common tasks. And most report that they received effective training to effectively implement common tasks.

[4] This notification was accelerated in the 2016-17 school year so teachers were notified of this expectation from the outset.

Because the tasks are developed by teachers, their familiarity with the assessment is better than it could be for a less familiar testing event. Teachers who do not participate in collaborative task development sessions have access to the task materials on the LibGuide and can consult with their Lead Teachers or Content Leads. Teacher support for administering the PACE assessment includes the online LibGuides, where tasks, rubrics, the Implementation Guidelines Manual, and other materials reside.

Teachers also support each other within schools and districts and outside their districts as well. They meet both formally and informally, in person and through shared working documents. They are also supported by their school administration, PACE District Leads, PACE leadership, and several expert consultants.

It is important to examine both the common tasks and the local tasks when considering the fidelity with which tasks are implemented. The common tasks are collaboratively developed and have a suite of task-specific student-friendly instructions and more generic supports (e.g., administration guide). Differences in implementation during the pilot phase of each task are reviewed and the task materials are revised to clarify, as needed.

Local tasks are exactly that, and could be used by only a single teacher. There is little worry that the local tasks are not administered as intended, but they may not be as fully developed or as in-depth as the common tasks. They are almost certainly not as well documented. This does not mean they are not effective performance assessment tasks. During interviews, teachers reported that their locally developed tasks have improved with every year of implementation. Starting with the 2015-16 school year, PACE will be auditing one local task per competency for every course in each Tier 1 district. These will be used to document local task quality and to provide feedback to teachers. When we think about implementation fidelity, we must consider the information provided by the local tasks. Scores from the local tasks are combined with scores from the common tasks to determine students' overall score and achievement level. So, training and support must be sufficient to allow teachers to create their own local tasks and administer them in a manner that supports the validity of the inferences made from annual test results. During the course of the evaluation we observed several local tasks being administered and it was clear that most PACE teachers understood how to design local tasks to support those inferences and how to assess their students in a way that elicits reliable performance data.

Claim 3b. Implementing the performance assessments as intended enhances and extends desired instructional practices

Teachers across districts expressed that implementing performance tasks has had a positive impact on their instruction. They commonly mentioned that PACE has had a positive impact on increasing the depth of knowledge (DOK) at which they teach and gives them real-time feedback that they can use to make "on-the-spot" adjustments to their instruction to better meet the needs of their students. Preparing students for the PACE assessments requires high DOK lessons and opportunities for students to apply and extend the content they've learned independently.

Unlike most large-scale assessment systems, which are focused on the estimation of student and/or school performance, PACE is also intended to influence instructional practices. Impact on instruction for most assessments would fall under the heading of unintended negative consequences. PACE leadership is not overly concerned about teachers "teaching to the test." PACE, ideally, supports "testing to what is taught." While most accountability assessments drive

instruction to at least some extent, their influence on instruction would not be viewed as positive by many educators. The high stakes and comprehensive nature of end-of-year tests may cause teachers to superficially teach many topics, spend days or weeks reviewing previously taught content, and spend instructional time on testing strategies.

PACE represents a significant step toward true integration of curriculum, instruction, and assessment. Several teachers reported that participation in PACE led to a qualitative shift in the way they approached assessment in their everyday work. Where assessment was previously used to differentiate superficial levels of performance and tests were a combination of items with varying degrees of difficulty and obscurity, they now focus on providing evidence that students have or have not achieved competency within a given content topic. This leads to more effective critical thinking about best practice, remediation and extension activities, and more productive reflection on and revision of day-to-day lessons.

Claim 3c. Student engagement and student learning increases/deepens when performance assessments are implemented as intended

Much like the shift in focus for teachers, PACE also represents a shift for students. Typical assessments are primarily focused on estimating achievement. Students learn content prior to the tests and then demonstrate their learning through their performance on the tests. PACE certainly has similar aspects, but because of the integrated nature of the assessments, students learn while testing as well. PACE tasks often require multiple classes to complete and might involve several steps (e.g., reading a novel, discussing the characters and their motivations, then writing a response to a prompt related to the novel). Because of the integrated nature of PACE, testing and learning are not entirely separate components of a student's day.

Teachers report higher engagement for their students and deeper learning of the content, during PACE assessments and as a result of improvements in their instructional practice that they attribute to participating in PACE. The majority of students report that they would rather take a PACE assessment than an end-of-year comprehensive test like Smarter Balanced or the New England Comprehensive Assessment Program (NECAP) test. When the students who indicated they would rather take a mostly multiple-choice assessment were asked why, they typically responded that they liked having some chance of getting the answer correct, even if they did not know the content very well. They commented on guessing and using test-taking strategies that were of no help on the PACE tasks. Others indicated that they preferred not to write their answers, as that was more difficult for them than a more traditional multiple-choice test. The students endorsing PACE discussed how closely the tasks were linked to their curriculum and how strong a measure of their abilities the tasks were.

Interim Goal 4. Scores are Accurate and Reliable

We found considerable evidence that students' scores and annual determinations are accurate and reliable. Scorers were effectively trained and PACE tasks were double scored. The common task was used to calibrate among the districts and to evaluate scorer accuracy.

Claim 4a. Scorers are effectively trained

PACE tasks, local and common, are scored using teacher-developed rubrics. These rubrics describe student work at four levels of competency. The teachers strive to make the distinctions as clear and concrete as possible. Adjectival scales (e.g., poor, acceptable, good, very good) are not acceptable. When teachers discuss the rubrics during the development process, they focus on the distinctions between the score levels and how to judge when students' work would represent one level versus the other.

After field testing of the common tasks, teachers come together to discuss and revise the tasks and scoring rubrics. During this process, the teachers score students' work. If there are inconsistencies or if the rubric is too vague to categorize students reliably, they adjust the rubrics. They also discuss the effectiveness and accuracy of the rubrics. For example, if the teachers agree that a students' work is exemplary, but some idiosyncrasy of the rubric forces them to give a lower-than-deserved score, they can adjust the rubric to deal with the issue. Once the rubric has been finalized, all the districts can score the task consistently.

During scoring, scorers begin with calibration sessions. These occur within districts and allow scorers to come to a common understanding of the application of the rubrics. They select and use anchor papers, or papers with known scores, to help calibrate and as a reference during the scoring process. While many teachers reported that the scoring process was time consuming, they were confident in their ability to score accurately and consistently.

The majority of teachers reported on the survey that the scoring rubrics for the common tasks are sufficiently clear and detailed to ensure that separate scorers scoring the same student work arrive at the same score, and that the scoring resources available on the LibGuide effectively explain how to score student work on the common tasks. Local tasks do not receive as much scrutiny. Training for developing and scoring rubrics for local tasks comes largely from teachers' prior experiences and from their work with the common tasks. They build in good scoring practice for the local tasks based on their experience and training.

Electronic score files are sent for generalizability analyses to the National Center for the Improvement of Educational Assessments (Center for Assessment). The Center for Assessment conducts cross-district comparability analyses and uses a standards setting procedure to establish district-level cut scores. Individual task scores are not adjusted during this process. The cut scores impact the overall determinations, and consequently the proportion of students in each of the achievement levels (1-4). Results provided by the Center for Assessment indicate that the overall scoring consistency is quite high and that few adjustments are necessary to the initially set cut scores due to inconsistent scoring (either too lenient or too strict) within the districts, indicating effective training for the scoring of PACE tasks. This process ensures consistency of scoring across districts. It is also the way that scores are made comparable across years.

Claim 4b. Scorers attain successful rates of interrater agreement and reliability

A sample of student responses to the common tasks are drawn for consensus scoring. Scorers work with a partner to rescore several students' work. Scorers may not be from the same district as the students whose work they score. Subsequent to the consensus scoring meeting, the scores from the central scoring group are compared with the scores from the district. If there is poor agreement between the district results compared to the consensus scored results, the scores on the common tasks are adjusted to account for the discrepancy. If the differences between adjusted common task performance is substantially different from local task performance, it may also signal a district level scoring bias. If such a difference is discovered, scorers can be retrained on a district by district basis.

Within-district inter-rater reliability is computed by the Center for Assessment. They determine whether a teacher scores more leniently or strictly by comparing the teachers' scores on the common task to the consensus scores on that task. The index they use for this purpose is a "deviance" index, which describes how far from the consensus scored papers an individual teacher scores (averaged across students). Several flags for potentially inconsistent scoring

have been established, but scoring for 2015–16 was quite consistent. While there were minor differences between subjects and by district, scoring for PACE common tasks by teachers was largely verified as accurate and consistent during consensus scoring.

The Center for Assessment also computes within-district rater agreement statistics (e.g. % exact agreement, % adjacent agreement) and Cohen's Kappa statistics for a sample of the double-scored common tasks (Evans & Lyons, 2016). Pairs of raters had exact agreement rates of between approximately 60 and 85%. There were substantial differences by grade, subject, dimension, and by district, but nearly all districts achieved greater than 60% exact agreement rates across all grade subjects. Kappa statistics indicate moderate to substantial agreement of ratings across all grades and subjects as well.

Samples of local tasks are also double scored. Teachers examine the results, but formal reliability statistics are not monitored during active scoring. Students' scores on the local tasks represent their work over the course of the year. They might be compared with more typical end-of-unit test scores. Unlike typical end-of-unit tests, students receive rubrics along with their PACE task instructions. This allows them to self-monitor as they work. If, at the end of the task the teacher score is different from the students' expectations, they can discuss the differences with the teacher. This provides the teacher with a quality check on the rubric and gives the student an opportunity to understand how to interpret and use the rubric to achieve the score they desire. Also, parents noted that the rubrics provide information to facilitate a discussion with their children about their performance on the task. The rubrics provide clear expectations for the students who use them, which improves the validity of the scores. The feedback teachers receive from the double scoring and from their interactions with students helps improve their locally developed tasks and rubrics to achieve better reliability.

Contextual Factors

While there are several contextual factors influencing the quality of PACE implementation worth mentioning, the largest stems from implementing PACE at the district level. Interviews with teachers and administrators in multiple districts yielded several district-specific positive and negative experiences with PACE. Districts vary in their capacity, student populations, and in the expertise and experience of their staff members. Early adopters of competency-based education had a significant advantage in implementing PACE. They already had a bank of mostly suitable locally developed tasks and were familiar with the design of competency-based rubrics. Their students had largely become accustomed to the kinds of tasks PACE requires. Districts that joined later had to build the infrastructure necessary to implement PACE.

District size plays an important role in PACE implementation as well. There are distinct advantages and disadvantages associated with being in a larger or smaller district, and it is not clear which is better. Smaller districts typically have only one teacher per grade/subject. In some cases, there may be only one teacher per grade; in elementary school this teacher is responsible for ELA, mathematics, and science tasks. This means that all of the work associated with developing and administering the local tasks is concentrated among very few people. Smaller districts often have to solicit help from outside the district to conduct double scoring. In addition, the requirement to submit copies of sample student work can be challenging because the smaller districts have very few support staff.

Larger districts have more support staff and typically have same-grade/subject teachers who can work as teams within districts, or even within the same school. This does not always mean that the teachers in larger districts have less work, however. The more students in a school who

take a PACE assessment, the larger the effort required for scoring. A very small district might only have 10 students who complete a task. A larger district could have a few hundred students completing a task.

District size can also influence teacher buy-in. Some of the smaller districts are close knit teams of educators, all of whom are supportive of both PACE and one another. In some districts, there are a few educators who are resistant to the implementation of PACE. This can cause strife for those who are committed to implementing the program with fidelity. It can be especially challenging when teachers bring PACE training or information back to their schools. These resistant teachers can have a larger impact in a small district, but larger districts are more likely to encounter them.

District size also impacts teacher expertise in PACE tasks. In larger districts, a subset of teachers participates in task development, necessitating that they keep their colleagues informed of the rich discussions from which they benefited. In small districts, a lone teacher could conceivably participate in task development for ELA, mathematics, and science—requiring substantial time out of the classroom.

In addition to district size, there are other contextual factors that may influence the implementation of PACE. Previous experience with competency based education and development of performance tasks made the transition to PACE easier for some teachers. Most PACE districts had previously developed Quality Performance Assessments (QPAs) that are similar to the PACE tasks. Those QPAs are often the basis for the local tasks for PACE, and participation in the QPA Institutes is an expectation for TIER 2 PACE member districts. Another important contextual factor is the perceived value of participation in PACE leadership roles versus the time requirements, especially time out of class. At least one district decided not to have any teachers serve as Content Leads. This district was heavily involved in developing the early common tasks. Many of the teachers from this district drafted the initial text for common tasks. The district decided to pull back from their leadership role to preserve the teachers' time in the classroom. Some teachers in the district were pleased with the decision, while others were disappointed that their role in PACE had been reduced.

Some parents, teachers, and students commented on the way that PACE tasks are scored. PACE tasks typically use "conjunctive" scoring, where multiple components of a task are scored separately, and the lowest of those scores becomes the final score. While conjunctive scoring is not a PACE-wide policy, common tasks are scored this way and teachers emulate the common tasks when they develop the local tasks. There was concern among multiple stakeholders that this method of scoring could result in lower than expected task-level scores. It was also not clear at what level the stakeholders were describing when discussing conjunctive scoring. Tasks might assess two or three dimensions within a subject area (e.g., mathematical modeling, computational accuracy, communication). The scoring rubric might contain several specific bullets describing students' performance under each dimension, each bullet scored 1-4. It is possible that conjunctive scoring could be done within a dimension, where the lowest bullet score determined the dimension score. This is the way the common tasks are scored. The dimension scores might be aggregated in some other way (e.g., by averaging). It is also possible that the lowest dimension-level score could be used to determine the overall task score. Parent focus groups in two districts referenced the scoring method as one reason that scores were lower than expected for otherwise high performing students, but they were not sufficiently specific to allow for fine distinctions regarding the scoring mechanism. Parents' concerns regarding task-level scoring was emphasized, in part, because a single task may be a large part of a student's 6-weeks grade (tasks are used for in-class grading as well as for the

annual determination), and because the student might only have a few task scores to contribute to the annual determination.

Scoring at the dimension level using conjunctive rules should result in better scorer consistency. If the scorers examine the evidence for each bullet, which can be very specific in terms of what the student is expected to produce, and then take the lowest scored bullet as the dimension score, we would expect a high degree of comparability at the dimension level (scorers only rate discrete well-defined evidence). Scorers might vary on other bullets, but if they agree on the lowest bullet, the overall task score would be the same. If the scorers use a compensatory approach, they must contend with defining "good enough" across the bullets.

Negative Consequences Minimized

PACE was implemented, in part, to reduce perceived negative consequences associated with large-scale, end-of-year standardized testing. PACE was designed to stave off reductions in the depth of learning of students, to promote critical thinking, and to integrate curriculum, instruction, and assessment into a cohesive system of education. We have discussed some of the ways that PACE has succeeded in reducing the negative consequences that already existed in New Hampshire schools, but it is also important to recognize potential negative consequences of PACE and to guard against them.

PACE tasks, especially science and English language arts tasks, can take a long time to implement. PACE tasks are designed to measure big, but reasonably discrete, ideas from the content standards. The developers must constantly ask themselves if the time investment to implement the performance assessments generates sufficient information to justify that time. Some of the science tasks can take more than a week's worth of classes to complete. Some of the English language arts tasks, because they may require that students read an entire novel in class, can take even longer. PACE task developers must guard against the tasks becoming so long that they unintentionally narrow the curriculum.

The PACE common task, in most ways, counts no more than any local task. It is used as an instrument to ensure scoring accuracy and reliability and as an equating tool to guard against cross-district incomparability. These additional uses, however, can cause teachers to give it much more attention than the local tasks. This added attention can be positive or negative. For example, one high school English language arts common task required students to respond to a text. One school chose a Shakespearean play as their text. The school then chose that same play as the winter drama production and staged the play with student actors for all the school prior to the administration of the task. While this is certainly not a prohibited activity, it may have given the school and its students an advantage over other schools that were not so savvy. Emphasizing the common task may limit the available time for other content. It may also create unintended differences between the way that the common versus local tasks are treated, which could, in turn, make attributions about scoring quality or other aspects of the PACE assessments less certain. Some teachers described spending a month or more in preparation for the common task (often including reading a novel aloud in class). If the task promotes strong lessons and broad and deep learning of the content, this level of effort may be entirely justified. If, however, the task represents a relatively discrete aspect of the overall curriculum, that time may be better spent. Interviews with PACE teachers revealed that most were very positive about the tasks and considered the preparation of students for them a major benefit. A small minority, however, indicated that the work on the part of teachers and students was disproportional to the benefit the students received.

PACE requires a tremendous amount of work on the part of teachers. While most teachers were very supportive of PACE, it was not uncommon for them to comment on the time and effort required to implement the program, including development of tasks and rubrics as well as task administration and scoring. Survey results indicate that approximately one fourth of respondents did not think that the time and effort required by the PACE initiative was worth the benefits Also, a few outlier responses obtained during interviews and focus groups suggested going back to Smarter Balanced. One goal of PACE is to generate enough tasks that development can become a more reasonable ongoing process of replenishing or revising only a few new tasks per year. Until that goal is reached, there is the potential for over-burdening teachers.

Once teachers develop units of study and associated performance tasks, they tend to use them for several years. The nature of PACE promotes this practice and, because of the complex nature of the tasks, we are not overly concerned with test security. There may be concerns, however, when the common task addresses the same, or closely related, content. Some teachers described having to abandon very strong units of study and local tasks because they were required to use the common tasks for that content. Using both would be time prohibitive and redundant, so they used the common task only. In one example, a teacher typically taught a life sciences unit on oceanography. The unit took advantage of the teacher's major area of study from college and was a highly-developed set of interconnected lessons for the students. However, because not all PACE districts have easy access to an ocean, the same content from her oceanography unit was tied to rivers and streams in the common task. She will likely teach the oceanography unit next year, when the common task changes, but she was disappointed that she had to replace it this year with a task she did not feel did as good a job of teaching the related content.

New Hampshire does not currently have a grade-by-grade curriculum for middle school science, but the common science tasks are grade specific. There is, therefore, some concern among educators that the tasks do not always match their curriculum. If, for example, one district teaches life sciences in grade 8, while another teaches physical sciences in grade 8, a common task in grade 8 related to life sciences could potentially disadvantage the latter district. The science tasks for middle school have been designed to address science and engineering principles and cross-cutting concepts, but these do not come content free. This issue may be resolved in one of two ways, based on current curriculum plans in New Hampshire. The New Hampshire Board of Education recently adopted the Next Generation Science Standards (NGSS). It is possible that this will lead to the adoption of a more consistent curriculum (at least by major content topic by year) in middle schools. PACE is also planning to allow districts to use matrixed/crossed designs with the common tasks. Alternatively, once a sufficient number of common tasks are developed, they can be administered based on the content of the course, rather than the grade of the student. So an eighth-grader taking physics might complete a physics task, while a sixth-grader taking a physics course might complete the same task. An eighth-grade in a different district might take a life sciences task. This would allow district-level control of curriculum, but may introduce new challenges for maintaining district-to-district comparability.

We conducted focus groups with a small number of parents in each Tier 1 district. While most of these parents were very supportive of PACE, there were a few who questioned the reliance of the program on performance tasks. This was especially true if the school or district adopted alternative reporting methods (e.g., changed report cards from traditional ABCDF grading to the 1-4 ratings for the PACE tasks). A few parents were concerned that colleges might not understand how the PACE tasks were scored and might inadvertently penalize their child because the grading system was so different from a traditional system.

Data Analyses

Comparison of Aggregate Data

Because PACE results are used in place of Smarter Balanced scores, it is important to consider the validity of PACE as an overall indicator of students' achievement in ELA, mathematics, and science at a specific grade level. This is the primary use of Smarter Balanced mathematics and ELA scores and we would expect PACE to provide similar results. We would not expect the results to be interchangeable. All of the differences in the design, purpose, development, administration, and scoring described earlier are expected to make PACE unique from Smarter Balanced. If the final results were the same, it would call into question if PACE truly represented a major shift in instruction and assessment.

New Hampshire does not require students to take both the Smarter Balanced and PACE assessments during the same year, so we could not directly compare assessment results for individual students. We compared the PACE results in aggregate to the Smarter Balanced results for the state, and also compared the results from 2015 to those from 2016. Analyses revealed that PACE and Smarter Balanced yielded differing results for classifying students as Proficient/Not Proficient, but those results were not so large or so variable as to call into question the similarity of the measured construct.

Student Level Correlation Results

In addition to examining scoring patterns across the PACE districts, we were also able to match a substantial portion of students' PACE scores from 2015 to their scores from 2016. PACE districts use differing scale scores, but use a common score level system (Levels 1-4), that has the same meaning for all PACE districts. We were able to correlate the scores across PACE assessments and across years to examine scoring patterns. Much like the comparisons of PACE and Smarter Balanced, we would not expect the correlations to be perfect, even for the same subject across years. If the correlations were perfect, we would not need to administer the assessments every year. Similarly, we expect scores across subjects to be correlated. Students who perform well in math tend to perform well in science and in ELA as well. So, we expect correlations that are strong and positive, but not perfect. This "Goldilocks" range of correlations that are neither too high nor too low indicate that the assessment system is functioning as expected.

We were also interested in patterns of correlations. Convergent validity coefficients (correlations between same subjects across years) should be higher than discriminant validity coefficients (correlations between differing subjects across years). We limited these comparisons to correlations across years because the time and instruction between assessments can attenuate correlations and we wanted to make the comparisons as similar as possible. All reported correlations for all grade pairs were statistically significant ($p < 0.01$).

Taken together, the correlation results provided strong evidence that PACE is functioning as intended. The correlations among the PACE subject areas within and across grades are similar to other statewide assessments (Dickinson & Thacker, 2009). Correlations within year among the PACE subjects were quite high, especially for elementary grades.

Recommendations

The recommendations in this section stem from the data collected during the course of the evaluation only. There is little literature that can be directly referenced and applied to a system like PACE. For that reason, there are no statements in the recommendations section that

reference aspects of similar successful programs. We did not find systems that were both successful and sufficiently similar to PACE to make direct comparisons.
The recommendations also reflect that PACE is currently functioning largely as intended. The early success of PACE is well documented in this summary report and in the associated technical report. No broad or sweeping recommendations are indicated. The recommendations included here call for additional monitoring or minor improvements to current processes. As the system expands, more substantial changes may become necessary, but this evaluation does not indicate a need for major modifications at this time.

Recommendations for Interim Goal 1: Stakeholders Are Committed to PACE

Recommendation 1: Monitor and Support District Engagement

PACE should regularly gauge local leadership support and target interventions when district leaders voice concerns or reduce their district's involvement with the program. PACE has done this for one district by helping support a PACE coordinator within the district with experienced consultants. As the program expands, these checks and interventions should become more routinized to ensure that all districts maintain adequate support for the educators implementing the program.

Recommendation 2: Evaluate Effectiveness of Collaboration Methods

PACE should evaluate the effectiveness of the new collaboration methods. While task development meetings with teachers from all Tier 1 districts were becoming unwieldy, one of the attributes teachers reported as positive was having direct input into the program. The more dispersed that input becomes, the less obvious individual teacher's input may be. If some teachers perceive the PACE program as coming from the outside rather than as a direct result of their own work, buy-in could suffer. Teachers regularly commented that the cross-district collaborations are a great source of professional development, and that they greatly value those opportunities. If the new collaboration methods reduce opportunities for cross-district collaborations, then teachers may perceive less personal value in PACE. Findings from the survey indicate that those teachers who had not participated in cross-district collaborations tended to have less favorable ratings of PACE. Regular monitoring and adjustments can help safeguard against this potential issue.

Recommendations for Interim Goal 2: Assessments Are Based on Sound Test Design Principles

Recommendation 3: Consider Additional Training/Supports for Teachers Not Directly Involved in Common Task Development

As the percentage of PACE participants directly involved in future common task development decreases (either through including a smaller number of teachers in a meeting or by expanding into additional districts), the professional development and training stemming from those activities may need to be supplemented with additional training. Teachers routinely reported that the process of developing the common tasks greatly improved their own task development process and their approach to assessment. As the program expands, it will be important to maintain that benefit for all participants.

Recommendation 4: Infuse Equity and Accommodations Training into PACE Activities

Include training on scaffolding and accommodations as part of the regular schedule of PACE activities. Despite quality documentation and training, teachers continued to report uncertainty regarding equity issues, especially for accommodating SWD. Scaffolding should be available to all students, including SWD, and is currently built into task development activities. The task instructions for teachers now include more information about appropriate scaffolding to ensure that all students can demonstrate their knowledge, skills, and abilities. These do not alter the nature or difficulty of the tasks, but provide entry into the various activities associated with it. Accommodations are provided to SWD as a means of improving access to the content of the task. They are based on students' needs and common supports identified for each student. They may alter the task, but should support measurement of the underlying construct. As the system expands and as attrition necessitates the inclusion of new teachers, it is important that these issues continue to be addressed to ensure both accessibility and validity.

Recommendation 5: Investigate the Impact of Reading/Writing Requirements on Accessibility

Investigate the impact of the reading and writing demands of the PACE tasks on accessibility and student performance. Several teachers indicated concerns that the reading and writing requirements for PACE were much higher than for traditional assessments. This can potentially result in reduced test score validity, especially for SWD. This phenomenon occurs when the reading/writing load interferes with the measurement of the intended construct. If, for instance, we are interested in knowing whether student understand and can perform computations associated with a mathematics concept, including a long reading passage to set up the task might interfere with a student demonstrating her math abilities. We recommend examining score patterns among the PACE tasks, course grades, and performance on comparison measures (e.g., Smarter Balanced) for students with and without disabilities as one way to investigate whether the reading and writing requirements may be impacting students' scores.

Recommendation 6: Routinize Timely Reviews of Local Performance Tasks

Evaluate the quality of the locally developed performance tasks and rubrics. As the pool of locally developed tasks expands, it is important to ensure that the tasks and rubrics are of sufficient quality to be used to generate student scores and annual determinations. Teachers report that their skill level in developing these tasks improves with each year of PACE participation, so it stands to reason that the validity and reliability of students' scores should improve with time. Instituting a system of regular task review will help ensure that happens. Some reviews have been completed at this time (by the New Hampshire Department of Education or by Stanford University), but teachers often reported that there was no feedback, or that feedback came very late from these reviews. Review of local tasks would benefit from a regularly scheduled, timely process.

Starting in the 2016-17 school year, districts will be required to submit one major assessment per competency per course, in addition to all local performance tasks in a common task template. At this stage in the evaluation, it is unknown how the assessments/tasks collected during the coming year will be reviewed, what feedback will be available to teachers/schools, or when that feedback will be provided. As this data becomes available, it will be very important to monitor the ways that feedback to teachers/schools is interpreted and used. This process has

the potential be very useful and positive for the PACE program, but it also has the potential to introduce unintended consequences.

Recommendations for Interim Goal 3: Performance Assessments Are Successfully Implemented

Recommendation 7: Plan for Future Research on the Impact of PACE on Teaching and Learning

The positive impacts of PACE on teaching and learning should continue to be externally verified beyond this evaluation. This may be part of a future research agenda when it becomes possible to evaluate the predictive strength of PACE results on college and career performance. In the interim, it may be possible to compare PACE versus non-PACE student performance on Smarter Balanced assessments, college entrance exams, or other measures.

Recommendation 8: Evaluate the Benefit of Time in Program on Outcomes

As the system expands, it may be possible to investigate the benefits of time in the program on instructional practice and student learning. If there is a benefit to spending several years in the PACE program, then that may bolster district-level support for the program and promote fidelity of implementation by educators. Teachers described a long period of adjustment and evolution of their teaching and assessment practices. It would not be surprising if there was a direct correlation between years in the program and benefits, both perceived and realized, on assessment practice and student learning. We would not expect this correlation to be perfect, however. Contextual factors such as district size, fidelity of implementation, and the effectiveness of district or school teams could certainly impact the effects of time in the program.

Recommendations for Interim Goal 4: Scores Are Accurate and Reliable

Recommendation 9: Consider Systematically Recycling Tasks

After the operational year, common tasks may still be used in place of, or in addition to, local tasks. PACE should consider some method of systematically repeating tasks across years as another check on the consistency of scoring. If tasks were repeated, previously scored “check sets” of student work from the prior year could be included in the current year. Score consistency across years could then be checked in a more systematic way.

Recommendation 10: Begin Tracking Performance from Year to Year

The PACE system has the potential for variability across years. Comparing performance across years will allow PACE to see where there are large changes in the proportions of students at each achievement level in any district and to investigate potential reasons for those changes. It is important to consider how changes in performance are reported and how they are characterized. Early reports to USED comparing student performance on PACE with performance on Smarter Balanced within and across years[5], as well as the data analyses completed for this evaluation, should be repeated annually. This will allow for continuous monitoring and by investigating anomalous results, PACE may be better able to identify potential threats to reliability and validity. Examples from this report include the lower correlations and reversed convergent/discriminant validity coefficient pattern for grades 7 and 8,

[5] See https://www.education.nh.gov/assessment-systems/documents/overview.pdf.

as well as larger than typical gains in math for grade 3. Conducting these analyses again next year will help PACE determine if these anomalies are random or if they represent some systematic difference in the way PACE is implemented by grade or subject.

We also recommend that PACE provide guidance for making valid inferences from annual performance information to schools, districts, and, if possible, the media.

End Goal: Students are College and Career Ready

Graduating students who are college and career ready is the ultimate goal of PACE. While we have found considerable evidence supporting the interim goals of PACE, it is still too early to evaluate college and career readiness. Once PACE has matured sufficiently and there are students who both experienced the PACE program and at least one year of college or career, we recommend that PACE support an ongoing research agenda to investigate claims under this ultimate goal.

Capturing the Story of PACE

PACE has lofty ambitions. Ideally, PACE will lead to an integrated competency based education system that is unbound by time in class, age, location where learning takes place, and other artificial methods of categorizing students. Instead, the system would focus on a core set of competencies and move students to the next phase of their education irrespective of when, where, or how the student achieves those competencies. The system will incorporate a large number of ways for students to demonstrate the competencies, and demonstration will take place in an on-demand way, where students can choose to complete a performance event (not necessarily limited to the current task format) when they are ready, rather than on a school calendar. Instruction would be more individualized and targeted toward the next competency the student needs to master. Such a system would reduce non-productive redundancy and allow students to learn at a much faster and more customized rate. Such a system would represent a dramatic shift from the traditional system of schooling.

PACE, as it is implemented currently, has taken steps toward this ideal. The PACE districts have begun identifying important competencies and they have designed performance tasks to measure those competencies. They have begun to build a bank of high-quality performance tasks that can be drawn on throughout a student's academic preparation. The have moved toward a more integrated system of curriculum, instruction, and assessment. Assessment is being woven into all aspects of teaching and learning, and the consideration of assessment when planning curricular sequence and planning lessons have increased among teachers since joining PACE. Students, even those who don't like PACE, describe the tasks as complex and difficult, but as strong measures of their knowledge, skills, and abilities.

The scores generated from the PACE tasks are sufficiently reliable for their intended use and they are valid for uses beyond those that can be gained from more traditional end-of-year tests. Students understand where they performed well and where they did not. Students can be given an opportunity to redo parts of their tasks once they have addressed the areas where they were not quite ready to demonstrate competency.

PACE has had a great deal of early success, but there is still a long road ahead if PACE is to realize all of its bold goals. First, PACE has to prove to be sustainable. The program is relatively new and a few highly motivated districts have been instrumental in implementing the system. As new districts join PACE, there will be challenges. Getting new staff members oriented to such a

complex new way of educating students takes considerable time and effort. If the experienced teachers train the new ones, they will need time to do so. They will need time in addition to the time they spend implementing PACE in their own schools and classrooms. There may also be performance gaps between the experienced and newly joined districts. These issues, as well as potential changes in the political and economic climate in which PACE is being implemented will likely challenge PACE. The sustainability of PACE will rely on demonstrating that the benefits of PACE continue to outweigh the challenges. For this to happen, PACE will require continuous feedback and improvement as the system expands.

The current PACE has been very responsive to challenges and has improved based on feedback. For example, task development and piloting have been accelerated to make sure every task is sufficiently piloted and revised before it is used operationally. Communication regarding data collection, in-person meetings, and other important calendar-specific activities has been improved and teachers have received this information earlier in the year. This helps teachers plan and makes the PACE system more readily implemented. PACE has begun to distribute minutes from Leads meetings as a means of ensuring common understanding of decisions and future plans. PACE has established Content Leads and Teacher Leads to limit the time teachers must spend outside their classrooms. All of these examples of program improvements resulted from PACE leadership responding to requests from teachers and/or feedback from this evaluation's interim reports.

In addition to the improvements PACE has already made, more enhancements are planned for the near future. PACE leadership plans to accelerate task development even more. The goal would be to allow pilot testing of the common tasks to begin in the fall semester if that is the most appropriate time in the curriculum to use them. This would allow a more genuine piloting of the tasks and provide data even earlier to facilitate review and revision of the tasks and rubrics. The PACE Content Leads are also discussing senior projects and senior exhibitions as a natural extension of this work. One of the monthly PACE Leads meetings was devoted to a presentation related to senior projects and exhibitions. The group decided that it was a useful idea to create a separate sub-group to meet monthly and explore ideas for implementing these new assessment components.

In addition to sustainability, PACE must also prove that it is scalable. New districts are joining PACE, but NH DOE recognizes the considerable challenges involved in scaling PACE statewide as it is currently conceived[6]. However, if PACE proves to be a substantially better system for educating students than the system that currently exists, it stands to reason that PACE should expand. PACE is currently adopted at the district level. This is, in part, because New Hampshire districts are extremely autonomous. It is, after all, the "Live Free or Die" state. Other states may not be structured similarly. Still, there is a great deal of preparation a district must do to become a Tier 1 PACE district. It would be difficult to suddenly implement PACE on a much broader scale because of the integrated nature of task development, teacher professional development, and collaboration. Getting a full state's population of teachers to suddenly begin to effectively collaborate seems unlikely. In New Hampshire, PACE began with a few highly motivated districts and is expanding carefully. This model seems to be effective for a system like PACE, and if the system is transported outside New Hampshire, other states may want to adopt a similar implementation plan.

[6] Indicated by NH DOE leadership and reiterated by district superintendents during interviews.

Formative Evaluation of New Hampshire's Performance Assessment of Competency Education (PACE) Final Report

Introduction

The New Hampshire (NH) Performance Assessment of Competency Education (PACE) is a program with ambitious and laudable goals. In spring 2015, the U.S. Department of Education granted New Hampshire a waiver from specific requirements of the No Child Left Behind Act and then the requirements of the Every Student Succeeds Act (ESSA) as part of a demonstration pilot program; a waiver extension was granted on October 6, 2016. In exchange, New Hampshire agreed to pilot an accountability system based on PACE. PACE is a competency-based approach in which students are provided meaningful opportunities to master and demonstrate critical knowledge and skills; those opportunities are embedded in instruction and occur at times appropriate to the curriculum timeline. In addition, PACE encourages educational change by providing an environment in which educators can improve their work rather than meeting the requirements of a traditional top-down accountability system.

The PACE program relies upon locally developed, locally administered performance assessment tasks aligned with local district grade and course competencies. These local competencies and local performance assessments are aligned to the State Model Competencies which, in turn, are aligned with national standards in each content area.

Participating NH districts administer Smarter Balanced assessments to grade 3 English Language Arts (ELA), grade 4 mathematics, and grade 8 ELA and mathematics, as well as the SAT to all grade 11 students. In addition to the local performance assessment tasks, a common performance assessment is administered in each grade and subject (ELA, mathematics, and science) without a state assessment. These common performance assessments were developed collaboratively by Tier 1 PACE districts and are administered in all Tier 1 PACE districts to ensure comparable evaluations of student performance across participating districts.[7]

The Human Resources Research Organization (HumRRO) was awarded a contract to conduct a formative evaluation of the PACE program in the Tier 1 districts between April 2016 and February 2017. The primary goal for this evaluation is to ensure that the PACE Leadership team has useful information to make decisions that advance the program's goals. This final report follows three interim reports. The first interim report described the development of the theory of action, development of data collection instruments, and observation of task development sessions (Becker, Thacker & Sinclair, 2016). The second interim report described interviews with eight PACE District Leads; site visits to three school districts; and observations of content expert training, the PACE Tier 1 Summer Institute, PACE Design Studio for Tier 2 and 3

[7] PACE districts are classified into three tiers: Tier 3 districts may or may not have written competencies, do not implement competencies at the classroom level with students, and may have no background experience with performance assessments. Tier 2 districts have course level and schoolwide competencies in place, have implemented competencies and competency-based education (CBE) learning in classrooms to some degree, and have limited experience with task-based performance assessments. Tier 1 districts have (a) implemented local competencies in schoolwide/classroom settings, (b) experience with performance assessments in a competency-based learning environment, (c) evidence of a commitment to transitioning to implementing performance assessment of competencies for accountability purposes for grades K-12, and (d) articulated an initial plan for accomplishing that transition.

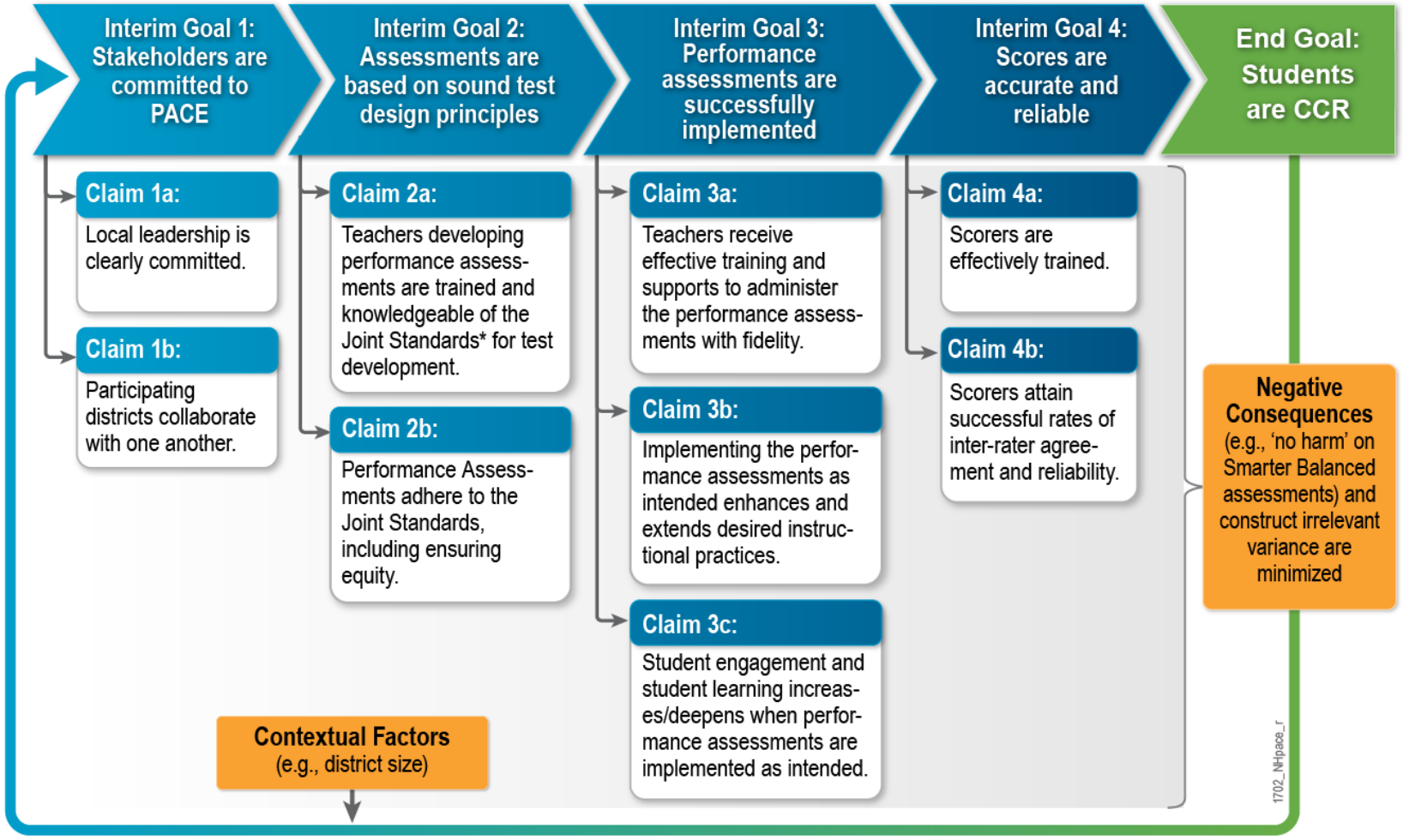
schools, and The New Hampshire Department of Education's (NH DOE) Annual Educator Summit (Becker, Thacker, Sinclair, Wiley, Woods & Dickinson, 2016). The third interim report described activities conducted between September and November 2016, including observation the September 2016 task development meeting, an interview with the PACE District Lead of a new Tier 1 PACE district, site visits to districts and schools, observations of PACE Leads meetings, a review of PACE standard setting and scoring, and preparation of a teacher survey (Becker, Thacker, Dickinson, & Sinclair, 2016). Each report built upon the prior one, identifying accumulating evidence regarding the goals of the PACE program. Each interim report captured the state of the PACE program at the time, with descriptions of what worked well, along with specific feedback intended to help PACE Leadership make continuous process improvements. This final report details data collection activities conducted between December 2016 and January 2017 and summarizes the evaluation as a whole. Throughout this series of reports we acknowledged when we observe incremental improvements that address issues we have raised. The *Data Collection* and *Results* sections of this report describe the activities conducted between December 2016 and January 2017. The *Summary of Findings by Evaluation Goals* and the *Conclusions and Recommendations* sections summarize the entire evaluation.

HumRRO's formative evaluation is focused on Tier 1 districts. As of the 2016–17 school year, NH PACE included nine Tier 1 districts. Four districts joined PACE in 2014–15: Epping School District School Administrative Unit (SAU) 14, Rochester School District (SAU 54), Sanborn Regional District (SAU 17), and Souhegan School District (SAU 39). A second wave of districts became PACE Tier 1 districts in 2015–16: Concord School District (SAU 8), Monroe School District (SAU 77), Pittsfield School District (SAU 51), and Seacoast Charter School (SAU 46). In addition, White Mountains (SAU 35) joined as a Tier 1 district in the 2016–17 school year. SAU 35 was included in limited fall/winter 2016 evaluation activities, such as an interview with the PACE Lead and the online teacher survey.

Our evaluation was guided by the PACE theory of action (TOA), which, per the Request for Proposal, reads: "*If we believe that all students must be college- and career-ready, then our system must advance students as they demonstrate mastery of knowledge, skills, and work study practices, which requires a comprehensive system of educator and school supports."* To flesh out this overarching TOA, HumRRO mapped the following nine success criteria that NH DOE provided to the U.S. Department of Education (US DOE) as part of its waiver agreement onto the TOA framework:

- Success Criterion 1: Gaining clear commitment from local leadership
- Success Criterion 2: Building cross-district leadership and cross-district collaboration
- Success Criterion 3: Developing high-quality performance assessments
- Success Criterion 4: Successfully implementing common performance assessments
- Success Criterion 5: Providing training and calibration
- Success Criterion 6: Reaching successful rates of inter-rater agreement
- Success Criterion 7: Producing "comparable" annual determinations
- Success Criterion 8: "No harm" on the Smarter Balanced assessments
- Success Criterion 9: Ensuring equity

The TOA, as expanded for our evaluation, is presented in Figure 1. In an argument-based approach to validation (Kane, 2013), the claims and assumptions that underlie each interim goal must be substantiated to achieve that goal. Lack of support for any one interim goal undermines all subsequent goals. For example, if the performance assessments are not administered as intended (i.e., Interim Goal 3), then the validity of the scores is called into question (i.e., Interim Goal 4), regardless of how high inter-rater agreement and inter-rater reliability are among the scorers. HumRRO drafted Figure 1 and the NH DOE approved this TOA in the project kickoff meeting.



** We understand that the PACE stakeholders are not test design experts and, therefore, that the AERA, APA, & NCME Standards are not firsthand knowledge for this audience. Consequently, our discussion with these stakeholders referred more generally to "high-quality assessment."*

Figure 1. PACE Theory of Action (revised after PACE District Lead interviews).

Data Collection

HumRRO's June, September, and Formative Evaluation: Final Reports included findings from (a) our observation of June 2016 Task Development Sessions for English Language Arts (ELA), mathematics, and science, (b) interviews with nine PACE District Leads, (c) site visits to four PACE districts, (d) observations of several PACE events: Training Content Experts, PACE Design Studio for Tier 2 and Tier 3 schools, PACE Tier 1 Summer Institute, NH DOE's Annual Educator Summit, September Task Development Meetings, (e) attendance at PACE Leads meetings, and (f) a review of PACE standard setting and scoring processes. The December 2016 report also described development of a teacher survey.

In this report we describe several data collection activities conducted subsequent to our December 2016 report: school visits in six Tier 1 districts, observation of December Task Development Sessions, attendance in a PACE Leads meeting, and results of the teacher survey.

Site Visits to Districts and Schools

Teams of two HumRRO staff conducted visits to six PACE Tier 1 districts in November–December 2016. In two cases (Sanborn Regional and Souhegan) this was a follow-up visit to supplement activities conducted during HumRRO's spring 2016 visit. Four districts (Concord, Epping, Pittsfield, and Seacoast School Districts) were visited only in fall/winter 2016.

Similar to the spring 2016 site visits, we endeavored to conduct the following data collection activities at schools serving each grade span (elementary, middle, and high) and across the content areas of ELA, mathematics, and science:

- Observe administration of PACE common performance assessments.
- Conduct focus groups with teachers administering PACE common performance assessments.
- Conduct interviews/focus groups with school administrators.
- Conduct student focus groups in grades in which PACE common performance assessments are administered.
- Conduct parent focus groups (one per district).

In each district, a district contact person worked with a HumRRO staff member to develop a schedule of activities for the visit.

Concord School District (SAU 8)

Two HumRRO staff visited Concord School District on December 6–7, 2016. The team attended Broken Ground Elementary and Rundlett Middle Schools on the first day and Concord High School on the second day. At each school, the team conducted student, teacher, and administrator focus groups. In addition, the visit included classroom observations at the Elementary School, an interview with the Assistant Superintendent, and a district-wide parent focus group.

Epping School District (SAU 14)

Two HumRRO staff visited Epping School District on December 1–2, 2016. This visit included an interview with the Superintendent; interviews with elementary, middle, and high school principals; a series of focus groups with teachers in grades 3 through high school; a series of focus groups with elementary, middle, and high school students; a parent focus group; and observation of an Arts class.

Pittsfield School District (SAU 51)

Two HumRRO staff visited Pittsfield School District on November 28–30, 2016. One day was spent at Pittsfield Elementary School and included a tour of the facility, observations of three classes, a student focus group, a teacher focus group, and a parent focus group. The other day was spent at Pittsfield Middle High School and included observations of three classes, a student focus group, a teacher focus group, a parent focus group, and an interview with the Dean of Instruction.

Seacoast Charter School (SAU 46)

Two HumRRO staff visited Seacoast Charter School on December 8, 2016. The visit included a school tour, a parent focus group, a teacher focus group, a student focus group, six brief classroom observations, and an interview with the Head of School.

Sanborn Regional School District (SAU 17)

Two HumRRO staff visited Sanborn Regional School District on November 29, 2016. This was a follow-up visit to supplement activities conducted during HumRRO's spring 2016 visit. The visit included several classroom observations at the elementary, middle, and high schools—including observation of a PACE task—and a parent focus group.

Souhegan School District (SAU 39)

Two HumRRO staff visited Souhegan High School on December 9, 2016. This was a follow-up visit to supplement activities conducted during HumRRO's spring 2016 visit. The visit included a parent focus group and an observation of students participating in a PACE science operational task.

Observation of NH PACE Task Development Meetings

HumRRO staff attended the PACE task development meetings on December 7–9, 2016. The task development meetings provided an opportunity for PACE Content Leads and Teacher Representatives from each Tier 1 district to come together to continue development of new PACE Common Performance tasks and to make evidence-based edits to PACE Common performance tasks that had been piloted in some schools.

Observation of PACE Leads Meeting

HumRRO staff telephonically participated in the January 2017 PACE Leads meetings. The PACE Leads meetings are held at the NH DOE offices and provide an opportunity for dissemination of information by PACE Leadership and discussions as a group.

Review of Annual Determinations

Because PACE results are used in place of Smarter Balanced scores, it is important to consider the validity of PACE as an overall indicator of students' achievement in ELA, mathematics, and science at a specific grade level. We analyzed extant data to compare the PACE results in aggregate to the Smarter Balanced results for the state as well as results from 2015 to those from 2016. We also correlated the scores across PACE assessments and across years to examine scoring patterns for evidence of convergent and discriminant validity.

Teacher Survey

HumRRO's evaluation included an on-line teacher survey. Plans for the on-line survey were discussed during the October 2016 Monthly Meeting with the NH PACE Leadership team and CIE. During the October 2016 Monthly Meeting the decision was made to survey all teachers in PACE Tier 1 districts,[8] including teachers responsible for administering NH PACE common performance task(s) and "other teachers" (i.e., those teachers at the school not administering NH PACE Performance Tasks). The decision was made to survey other teachers in order to more broadly investigate the schoolwide impact of PACE. An initial branching item on the survey identified those teachers not administering NH PACE common performance tasks and routed them to an abbreviated version of the survey that contained a subset of items that also appeared on the survey for teachers administering NH PACE common performance tasks (to facilitate comparisons in responses). From this point forward, the version of the survey for teachers administering NH PACE common performance tasks is referred to as the PACE Teachers Survey and the version of the survey for teachers not administering NH PACE common performance tasks is referred to as the Other Teachers Survey.

NH DOE provided HumRRO with a list of individual email addresses for all teachers in each Tier 1 district and HumRRO emailed a unique URL link to individual teachers. HumRRO emailed reminders to non-respondents at key points during the survey window. HumRRO removed teacher email addresses from the final data set prior to analysis to ensure anonymity of responses.

The draft content for the survey items was delivered to the PACE Leadership team and the University of Kentucky's National Center for Innovation in Education (CIE) on November 2, 2016, for their review and feedback. The survey was designed to minimize respondent burden as much as feasible. The survey included several selected response items and one open-ended item. The survey items were developed to address the nine success criteria described in the Introduction of this report. In addition, survey items were developed to capture information about the effectiveness of various changes that were recently implemented by the PACE Leadership team (e.g., addition of Content Leads). Collecting teacher feedback on the impact of such changes helped ensure that formative information for the PACE Leadership team and CIE was captured.

Feedback was received from both the PACE Leadership team and CIE on the draft survey items. HumRRO incorporated that feedback, and on November 16, 2016, HumRRO provided a

[8] This includes White Mountains, which joined the Tier 1 districts in the 2016-17 academic year. Given that teachers from this district were new to PACE, a disaggregated analysis of responses from White Mountains was conducted to determine whether their responses would skew results if included in the overall results.

Microsoft Word document with the survey revisions noted via track changes, along with URL links to the pilot versions of the on-line surveys. On November 22, 2016, HumRRO received final approval from the PACE Leadership team on the survey items. The final set of survey items (in Microsoft Word format) is provided in Appendix A for PACE Teachers and Appendix B for Other Teachers.

HumRRO provided NH DOE with a brief written description of the survey and the rationale for collecting email addresses for all teachers in Tier 1 districts. NH DOE shared this information with PACE District Leads, gathered teacher email addresses from the nine Tier 1 districts, and forwarded them to HumRRO between November 10 and November 18, 2016. HumRRO standardized the file formats and resolved discrepant records.

Prior to the survey launch, HumRRO emailed the PACE District Leads to inform them that the emails with the link to the survey would be distributed to their teachers on November 28, 2016. The survey launch was accomplished on schedule, with the exception of the Rochester School District. Due to an especially tight spam filter, the emails with the link to the survey were not successfully delivered to the teachers in the Rochester School District until December 8, 2016. The response deadline for the survey was initially scheduled for December 16th; however, given the delay in delivery of the survey to the teachers in the Rochester School District the survey deadline was extended to December 23, 2016. This same extension was provided to all Tier 1 districts. Weekly reminder emails were sent to all non-respondents and to all individuals who had accessed the survey link, but not completed all the items.

Results

The following sections describe observations and findings from each data collection activity described above. While the PACE system includes both common and local performance tasks, this section focuses more heavily on the common tasks, as requested by PACE Leadership during the kick-off meeting.

Site Visits to Districts and Schools

Concord School District (SAU 8)

Overall, the attitude of the district is positive towards PACE tasks. After visiting the three schools, a number of overall themes were noted and are summarized below:

- Concord has a strong administration that is invested in PACE. They work with additional organizations to prepare the district for task readiness.
- The administrators and teachers praised the PACE program for professional development, and collaboration among and between districts. Teachers particularly enjoyed sharing ideas about learning in the cross-district collaborations.
- Administrators and teachers acknowledged the large amount of time required for PACE activities, but felt it was a necessary component of the PACE journey. They felt PACE brings their district and classrooms to a higher level of instruction.
- The assistant Superintendent spoke in favor of PACE in her district. She indicated the district dedicates considerable amounts of time for teacher involvement in the program, and the growth of teachers and students is favorable. She indicated the district continues to grow in its knowledge and implementation of PACE.

- Teachers stated there is a high level of buy-in from staff. They felt it was higher for those most involved with PACE activities.

- Teachers who did not participate in in PACE collaborative activities, such as task development, were less clear on the level of expectation required of them.

- Teachers explained PACE assessments are designed to occur organically after what is taught in the curriculum. They felt that PACE was a better fit for their students and a more productive use of time than traditional standardized assessments. PACE challenges the students on what they learned, not just how to take a test. Since PACE tasks contribute to a student grade, teachers discussed the challenge of conducting makeup sessions if students are absent during the assessment. Students remembered past PACE operational tasks as engaging, relevant, and a natural extension of classroom learning. They felt PACE allowed them to express themselves and personalize the task. They remarked they appreciated the different style of testing.

- Many PACE common tasks include group activities, culminating in individual write-ups that are used for scoring. Some students remarked that it was challenging at times to work in a group setting. However, they also noted this was positive because it enforced real world, collaboration skills. Other students remarked on the difficulty of group work with students not as motivated to do well, but also felt this was mitigated with individual work. Some students remarked they would prefer a choice of solo work or the ability to choose who participated in their group.

- Younger students commented that some parts of the rubrics were confusing or too complex. Older students stated the rubrics were similar to those they use for classroom assignments.

- Most parents had received little communication about PACE and were generally unaware of PACE tasks being administered, so they advocated for increased communication. The Superintendent explained that district leadership was hesitant to distribute information widely during the early stages in which staff were gaining familiarity with the program.

- Parents appreciated PACE's focus on critical thinking skills applicable to the real-world. They noted that PACE results in their children being more confident and having increased opportunities to master learning.

Epping School District (SAU 14)

A summary of the key themes that arose from our visit to Epping School District is presented here. Overall, the findings indicated that reactions to PACE were largely positive. The themes included:

- The Superintendent and others noted that she has lost teachers because of PACE. She explained that PACE is a small piece of the Epping vision and PACE takes time from teaching.

- The Superintendent commented that reporting is a challenge. Schools may be reporting scores on multiple scales (e.g., 1-100, letter grades, 1-4 competency ratings) which can lead to confusion.

- Epping has a part-time (60%) consultant helping with many PACE activities. She offloads work from the teachers, facilitates communication, and serves as a resource for teachers. Her contributions were praised throughout the visit.

- Teachers cited the challenges of time and communication. Teachers and administrators noted that communication and organization have improved, citing the accelerated establishment of a calendar of activities this year.

- Three teachers represent Epping at collaboration meetings and none were identified as Content Leads. Various teachers expressed frustration and feeling "out of the loop."

- Some teachers credited PACE with improvements in their classroom assessments, although various teachers noted that they had to drop some lessons to make time for PACE common tasks.

- Several teachers expressed frustration with the amount of unpaid work required to participate in PACE. Activities ranged from scanning and copying student work (with names covered), scoring, and task development activities.

- Teacher and administrator opinions differed as to whether PACE is more accessible than traditional standardized tests.

- Parents who attended a focus group were familiar with PACE tasks and raised concerns about PACE scoring (e.g., due to the conjunctive scoring an "A" student earned a "1" on a PACE task), potential for students to copy other students' work (e.g., solar cooker project), scores sent home without teachers' knowledge, perception that rubrics are poor. On the other hand, parents recalled their children's enjoyment of specific tasks.

- Most students who participated in focus groups enjoyed the PACE tasks and preferred them to multiple choice exams. One student commented that it was "good to get my brain working a little bit." A few exceptions at each grade level preferred multiple choice exams because they don't require writing, and the student can use the process of elimination to determine the answer on a multiple choice test. One high school student, on the other hand, commented that the right/wrong nature of multiple choice tests means you cannot learn much about what you did wrong, while on PACE you can get partial credit and "you might learn something new."

- Students had a high awareness of their teachers' perceptions of PACE tasks. If a teacher expressed dissatisfaction with a PACE task or rubric, the students seemed to share the opinion.

- High school students expressed a common theme that the PACE questions and instructions were more complex than they were used to.

- In a classroom observation of a high school arts class, students debriefed their performance task using a rubric. Students indicated that this was more interactive than a typical lesson and reported that the task was worthwhile and would work well for other art media.

Pittsfield School District (SAU 51)

A summary of the key themes that arose from our visit to Pittsfield School District is presented here. Overall, the findings indicated that reactions to PACE were mixed.

- Student opinions varied regarding the difficulty/ease of the tasks.
- While some students commented that it was frustrating to have to wait for the rest of the class to catch up before proceeding to each new part of a task, they generally agreed that the PACE tasks were challenging and preferable to repeating the "same kinds of problems over and over."
- Students were familiar with the rubrics and used them to ensure they completed all parts of a task. They expressed irritation that the conjunctive scoring meant that they received the lowest partial score as their overall score for a task.
- Only one student repeated a task to improve his class score, although most students were aware they had this option.
- Although students claimed not to have learned anything during one of the mathematics tasks, when asked a parallel problem by the focus group administrator, the students were able to apply what they learned.
- Teachers noted that PACE was mandated by their district so initial buy-in was weak. However, once teachers completed the entire PACE process, buy-in increased substantially. One teacher commented that "It's hard to justify just having kids do a multiple choice test on the computer when you see how rich the [PACE] results can be."
- One teacher noted that in a multi-age classroom of students in grades 4 and 5, five common PACE tasks were administered. The teacher commented on the need to organize and keep up with many details to accomplish this.
- Some teachers expressed that training was a challenge and they were left to access the website and "figure it out." However, teachers noted that the LibGuide provided valuable information.
- While teachers praised the performance tasks, they also noted that the data collection requirements were a burden. In the 2015–16 school year, teachers did not learn of the data collection requirements until well into the school year, necessitating some backtracking and extra work. While these requirements were identified earlier for the 2016–17 school year, the amount of work remained a significant challenge. These teachers split the burden among themselves to acquire the sample of student work for 18 students per grade. They felt comfortable identifying high, medium, and low performers after about 3 units.
- Teachers described various degrees of revising their teaching to accommodate the PACE common tasks. These ranged from adding language and tools (such as sketching) to everyday instruction so students were familiar with the format, to re-organizing the schedule to ensure relevant lessons and competencies were presented in time for the PACE task. In the most extreme case, materials were accelerated to Algebra 1 from an Algebra 2 course. One teacher voiced frustration that now they are teaching to the test, both in format and content; other teachers did not express support for this opinion.
- Some teachers allowed students to repeat the performance task to improve their classroom grade while other teachers did not.
- Teachers noted that supporting students not on reading level was extensive due to the heavy reading load of the performance tasks.

- Teachers expressed worry about cross-district difference in the standardization of the PACE administration. One teacher suggested, and others agreed, that training videos showing correct task administration would be helpful to ensure all teachers are on the same page.
- Parents expressed that because performance tasks are the common way of assessing at this school, the PACE task "becomes just another day."
- Parents praised the act of familiarizing students with the rubrics in advance so they know expectations ahead of time. They opined that whether students are surprised by their scores or not depends on the teacher making the rubric clear.
- One parent commented that in some ways the scoring of performance tasks is more arbitrary, not clearly right or wrong like a multiple choice test.
- Parents disagreed as to whether students were more engaged in PACE tasks than other classroom assignments or tests. Some felt that the application piece makes the learning deeper, while others suggested that the tasks may be more entertaining and interesting but not necessarily deeper.
- Multiple parents noted that the PACE tasks train students to self-critique and have ownership of their learning. The school held student-led conferences that the parents found to be very positive.
- Most observed class lessons were interactive, including students working out mathematics problems on the whiteboard; class-wide editing to make sentences more specific before revising their own writings and sharing with partners; and two classes in which the teacher demonstrated, followed by small groups conducting an experiment. Two classes did mostly independent work.

Seacoast Charter School (SAU 46)

A summary of the key themes that arose from our visit to Seacoast Charter School is presented here. Overall, the findings indicated that reactions to PACE were largely positive. The themes of support for PACE that arose included:

- Parents indicated they prefer PACE to more traditional testing formats.
- Teachers indicated that PACE has had a positive impact on their instructional practice. Among the impacts mentioned were increased awareness and depth of knowledge (DOK), more meaningful discussions and collaborations, a keener focus on the content that students need to be exposed to, and an enhanced understanding of the multiple ways that students may demonstrate learning.
- Teachers and the school administrator indicated there is strong "buy-in" among staff to the PACE philosophy.
- Teachers and the school administrator found opportunities for collaboration with other districts through PACE scoring and body of work ratings to be very useful.
- Students generally reported fond recollections of past PACE operational tasks and characterized them as fun, relevant, and a seamless part of their overall learning experience.

- The school administrator indicated that PACE allows teachers more instructional time due to less time needed for test preparation activities.

The following challenge was consistently mentioned across respondents:

- Teachers and the school administrator cited limited time and resources as an ongoing challenge, particularly for a small school such as theirs.

Sanborn Regional School District (SAU 17)

This follow-up visit included several classroom observations and a parent focus group. Classroom observations included three grade 4 social studies classes developing brochures about New Hampshire and three grade 4 classes studying mythology. Varied approaches were employed in the classrooms, such as reading a play aloud with each student assigned a character in the play; answering prompt questions; and creating products to compare two stories using posters, Venn diagrams, or tables. We observed grade 5 students taking various self-directed approaches to building a creature with wildly varied features to accomplish specific tasks through drawing, writing by hand, or preparing a response on the computer; grade 8 science students beginning a multi-session exercise in building a trebuchet; and grade 10 students taking an ELA writing assessment.

Key themes included:

- Students were highly engaged in their tasks and, when asked, could clearly communicate what they were doing and why.
- Some tasks were group-based and others were individualized; students appeared comfortable with both formats.
- A few technological issues were dealt with promptly.
- Parents expressed familiarity with PACE, competency based education, and several specific PACE tasks. They were enthusiastic about all of those educational experiences.
- Some of these parents had older children or their children transferred from another school, so they were able to compare educational experiences with and without PACE. One parent was “blown away by the project work here” and explained that children’s enthusiasm improved dramatically.
- One parent noted that lots of folks do not understand this new approach to education, and that any time there is a paradigm shift the community takes a while to get up to speed. The principal provided “lots of reports” to keep interested parties up to date. Some parents were specifically concerned that colleges might not understand the grades/class standings employed under PACE, but the principal alleviated concerns by explaining that colleges—especially Ivy League institutions—prefer this kind of reporting. Another parent noted that adults without children sometimes express these concerns without sufficient information.
- One parent recently had an opportunity to hear about the positive reputation of the district in the state and in the country, and noted that local parents typically do not know about this. Other parents agreed, noting that people in the community sometimes get partial information or they misunderstand and spread rumors. They pointed out that the people who complain are typically those who do not attend meetings or seek information on the official website.

- Parents praised the implementation of PACE tasks as a typical unit test rather than as a big event "on the principal's calendar with big red circles." One parent of a child with an anxiety disorder observed that a PACE task does not produce the same level of anxiety as a test.
- Parents described the rubrics as sensible and easier to read than the competencies. Having a marked-up rubric allowed parents and children to have more meaningful conversations about expectations, including discussions about where the child succeeded or fell short.
- Parents noted that PACE tasks encourage a deeper level of understanding than a traditional multiple choice test. They said that the preparatory work, and the task itself, caused students to retain their learning longer.

Souhegan School District (SAU 39)

A summary of the key themes that arose from our follow-up visit to Souhegan High School is presented here. Overall, the findings indicate that reactions to PACE were largely positive. The themes of support for PACE that arose were as follows:

- Parents cited improved retention of learning and reduced test anxiety among the benefits that PACE had for students.
- Students participating in a PACE operational task were engaged and appeared to enjoy the experience.

Observation of NH PACE Task Development Meetings

The PACE Task Development meeting was held on three consecutive days from December 7–9, 2016. Educators from Tier 1 districts, along with PACE Leadership and other experts, convened for one day. Mathematics educators met one day; science educators met the next; and ELA educators met on the final day. Most of the educators had participated in previous task development meetings.

Mathematics

The mathematics task development meeting lasted a full day and began with a large-group presentation on Depth of Knowledge (DOK) by PACE leadership. After an overview of key DOK-related concepts, an example task was distributed among participants for discussion of the task's DOK level. Finally, recommendations were offered for considering DOK expectations during task development activities. Participants later reported finding this presentation extremely helpful to their subsequent task development work.

The remainder of the day was spent in breakout rooms organized by grade level/mathematics content area. Each breakout session began with a reflection exercise to assist in the establishment of group norms. These norms were posted at the front of each room to serve as a reminder of expectations around communication and collaboration during task development work. Participants appeared to take these norms to heart and maintained a collegial and collaborative environment throughout the day.

Breakout rooms varied slightly in their organization, depending on the task development activities to be completed. For example, the Grade 5 group consisted mainly of teachers who had not previously participated in test development activities, with two experienced PACE

Content Leads serving as facilitators. These two facilitators had drafted the initial version of the task during the September task development meeting, so the full group's focus was on editing that task and developing the initial scoring rubric. On the other hand, the Geometry room was split into two separate groups, with one group working on editing a draft new task and rubric and the other group making edits to another task/rubric based on student data collected during a pilot of that task. Regardless of the specific activities, participating teachers engaged in deep discussions about their expectations around student performance and how to design tasks and rubrics that will allow students to demonstrate performance at all levels.

All groups were observed making use of Google Docs to track and monitor changes to the tasks and rubrics. In one group, the rubric under development was divided into sections and teachers worked simultaneously in a single document, allowing for efficient use of time. In another group, teachers agreed to provide feedback in real time via Google Docs during the task pilot phase, so that any issues could be addressed quickly.

Science

Five teams worked in separate rooms to develop the science tasks: grade 4, grade 8, Physical and Earth, Biology, and Chemistry. Groups began by establishing group norms for communication and collaboration during task development work, then proceeded to work on pilot tasks at various stages in development.

Breakout groups varied somewhat in their approaches to task development. For example, the grade 4 group had 15 participants, and worked in small groups to complete separate parts of the task. In some groups, clear leaders directed the discussions while other groups were more diffused in their approaches. PACE Leadership provided targeted coaching during breaks to select Content Leads; the observer noted improvements in the group dynamics.

Some teachers had pilot tested early versions of the science performance tasks. They described the event to the groups, including implementation decisions not specified in the task but that were made by the teacher, aspects of the tasks that worked well and the aspects they found challenging, and student reactions to the tasks. In some cases, the teacher shared student work samples and the group used the rubric to score them. The discussions of the pilot experience were very useful to guiding refinement of the science performance tasks.

PACE Leadership visited the groups periodically to check on progress and address any questions. Questions included specific details of tasks and rubrics as well as whether a task could be administered to students in a different grade level than was intended. The group was told that during the operational year the tasks must be administrated at the target grades; in later years, schools are free to administer the tasks when they see fit.

Participants paid particular attention to ensuring the tasks and associated materials would be clear to teachers not participating in task development. For example, the grade 8 science group included a resource list for teachers who had not taught this topic before.

ELA

Teams developing ELA performance tasks for grades 4, 5, 6, 7, 9, and 10 worked in separate rooms. They followed procedures similar to the mathematics and science meetings by collaboratively developing norms at the outset, such as begin and end on time, stay on task,

assign a timekeeper, ensure everyone has a voice, keep an open mind, and rules for making group decisions, among others.

Activities were similar to those described above for the mathematics and science meetings. Some of the ELA groups held deep discussions about providing a checklist as part of the task, without being prescriptive and providing too much scaffolding. They also discussed the use of preparatory mini-lessons included with the PACE task materials. The grade 7 group decided that these mini-lessons were not required, but rather the teacher should consider using the mini-lessons to check that students possess the skills to succeed on the task. For example, if the teacher already addressed these skills earlier in the year, there would be no need to conduct the mini-lesson.

When available, groups scored the student work. None of the participants in the grade 6 group had piloted the task so they took the task themselves.

Some groups developed separate scoring rubrics for the teachers and students.

At the end of the day, the groups developed an agenda for the next meeting and included activities such as analyze student samples, refine the rubrics, and finalize the task prior to the start of the next school year.

Overall

Observers noted marked improvements in the focus, organization, and productivity of these fall 2016 groups relative to previously observed task development meetings. The next task development meetings will be in March 2017 and each breakout group determined which (and how many) teachers could pilot the task prior to that meeting. In the March meeting, the pilot tasks and supporting materials will be revised as needed, and volunteers will be recruited to pilot test the revised tasks before the end of the school year. Subsequent to the meeting, PACE Leadership determined that 98 teachers will pilot test the mathematics, science, and ELA PACE performance tasks across participating NH Tier 1 districts between January and March 2017.

Observation of PACE Leads Meeting

Monthly PACE Leads meetings were held at the NH DOE offices and provided an opportunity for dissemination of information by PACE Leadership and group discussions of topical issues. PACE Tier 1 districts were expected to provide a representative at these meetings even if the PACE Lead was unavailable. In addition to PACE Leadership and PACE Leads, additional parties were invited to present or contribute to discussions, as topics warrant. An agenda was distributed in advance of each meeting. Official meeting notes were not distributed; each PACE Lead was responsible for sharing information as appropriate in his or her district. In response to feedback from the field, in January 2017 PACE Leadership notified attendees that, for future meetings, official meeting minutes will be prepared and made available on the LibGuide. Over the course of the evaluation, HumRRO staff attended four of these meetings (three by phone).

Agenda topics routinely included updates on US DOE actions, PACE task development, and recent events. In addition, targeted topics were included for in-depth discussion, such as updates from the Center for Assessment, including a walk-though of PACE data reports, implications for improvement based on PACE data, and data privacy. Throughout these meetings, PACE Leadership encouraged active participation and input by the PACE Leads.

A current topic that was discussed in multiple meetings developed into its own series of meetings, scheduled in conjunction with the monthly PACE Leads meetings: senior exhibitions/senior projects. Rich, ongoing discussion of how senior projects (as well as activities in earlier grades) have already taken place in some districts. Additionally, information has been provided to other districts that are not ready for such a commitment so they can take initial steps to position themselves to implement senior projects over time. The senior exhibition/senior project has evolved into a natural extension of the performance tasks at the center of PACE.

Review of Annual Determinations

Comparison of Aggregate Data

Because PACE results are used in place of Smarter Balanced scores, it is important to consider the validity of PACE as an overall indicator of students' achievement in ELA, mathematics, and science at a specific grade level. This is the primary use of Smarter Balanced mathematics and ELA scores and we would expect PACE to provide similar results. We would not expect the results to be interchangeable. All of the differences in the design, purpose, development, administration, and scoring described earlier are expected to make PACE unique from Smarter Balanced. If the final results were the same, it would call into question if PACE truly represented a major shift in instruction and assessment.

New Hampshire does not require students to take both the Smarter Balanced and PACE assessments during the same year, so we can't directly compare assessment results for individual students. We can, however, compare the PACE results in aggregate to the Smarter Balanced results for the state. We can also compare the results from 2015 to those from 2016. Figure 2 presents this information for mathematics.

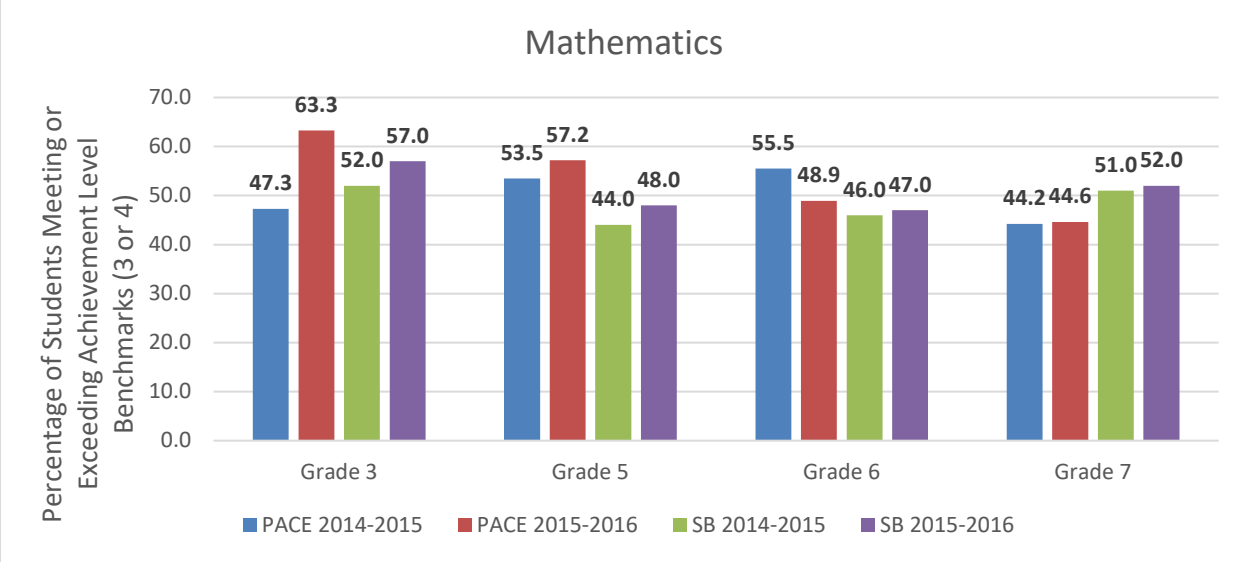


Figure 2. Comparisons of PACE and Smarter Balanced Mathematics Results for 2015 and 2016[9]

Figure 2 shows us that the PACE results tended to be somewhat higher than Smarter Balanced for grades 3, 5, and 6, but somewhat lower for grade 7 (except for grade 3, for which PACE

[9] PACE results were not available for high school for 2015. High school results are presented in the full technical report.

scores were lower in 2015 and higher in 2016). If we look across years, we see that the Smarter Balanced results improved from 2015 to 2016 in all grades, while PACE improved in grades 3, 5, and 7, but declined in grade 6. PACE results also tended to be more variable from year to year. The results are similar and indicate that PACE and Smarter Balanced tended to classify reasonably close to the same proportions of students as Level 3 or above.

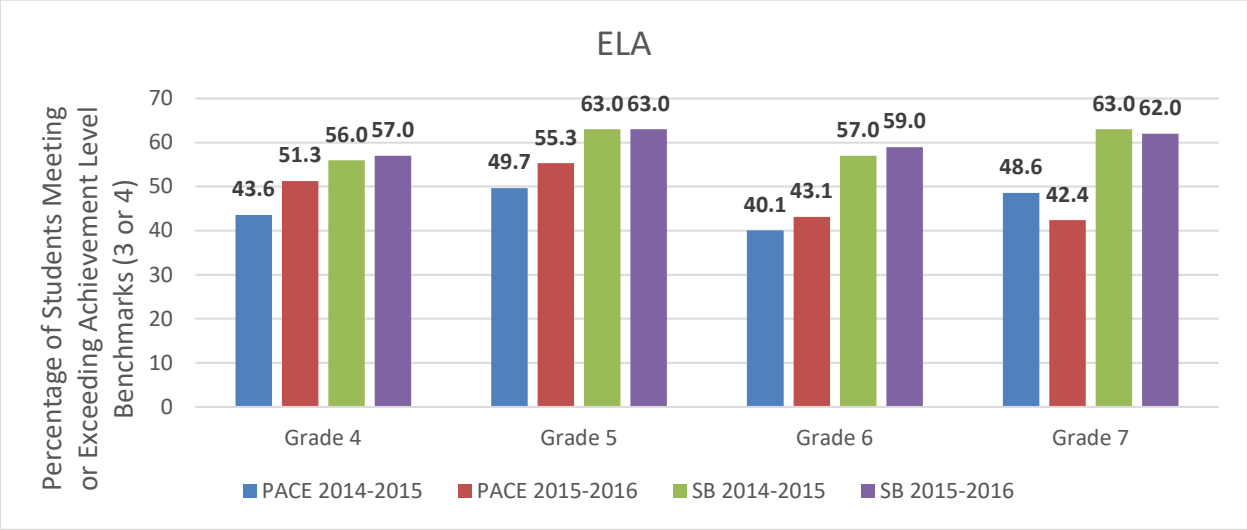


Figure 3. Comparisons of PACE and Smarter Balanced ELA Results for 2015 and 2016[10]

Figure 3 provides the same information for ELA. The results were even more consistent for ELA. In grades 4, 5, 6, and 7 PACE classified fewer students at Level 3 of higher than Smarter Balanced. Both assessments showed improved performance from 2015 to 2016 for grades 4, 5, and 6, while both assessments showed a decline in performance at grade 7. This demonstrates that PACE and Smarter Balanced were likely classifying students similarly for ELA. Taken together, Figures 2 and 3 indicate that PACE and Smarter Balanced yielded differing results for classifying students as Proficient/Not Proficient, but those results were not so large or so variable as to call into question the similarity of the measured construct.

Student Level Correlation Results

In addition to examining scoring patterns across the PACE districts, we were also able to match a substantial portion of students' PACE scores from 2015 to their scores from 2016. PACE districts use differing scale scores, but use a common score level system (Levels 1–4), that has the same meaning for all PACE districts. We were able to correlate the scores across PACE assessments and across years to examine scoring patterns. Much like the comparisons of PACE and Smarter Balanced, we would not expect the correlations to be perfect, even for the same subject across years. If the correlations were perfect, we would not need to administer the assessments every year. Similarly, we expect scores across subjects to be correlated. Students who perform well in math tend to perform well in science and in ELA as well. So, we expect correlations that are strong and positive, but not perfect. This "Goldilocks" range of correlations that are neither too high nor too low indicate that the assessment system is functioning as expected.

[10] PACE results were not available for high school for 2015. High school results are presented in the full technical report.

We are also interested in patterns of correlations. Convergent validity coefficients (correlations between same subjects across years) should be higher than discriminant validity coefficients (correlations between differing subjects across years). We limit these comparisons to correlations across years because the time and instruction between assessments can attenuate correlations and we want to make the comparisons as similar as possible. Tables 1–5 present the correlations among the PACE scores that were available for this evaluation by grade pairs. Each correlation represents between 386 and 455 students (the number that could be matched from the 2015 and 2016 data per grade pair). All correlations are for Achievement Levels (1–4) due to differences in scale scores by district. All reported correlations for all grade pairs were statistically significant ($p < 0.01$).

Table 1 presents the correlations for grade 3 in 2015 and grade 4 in 2016. Third-grade students only had mathematics scores for 2015, but had ELA and science scores in 2016. This yielded 3 correlations, all of which were for differing subjects. Not surprisingly, the highest correlation (0.637) was between ELA and Science, both administered in 2016. Given the potential attenuating effect of using Achievement Level instead of scale score, the correlations were strong and positive.

Table 1. Correlation Table Grade 3-4

	Math 2015	Science 2016
Science 2016	.487	
ELA 2016	.317	*.637*

Table 2 presents correlation results for grade 4 in 2015 matched to grade 5 in 2016. This table presents the first available convergent validity coefficient (ELA 2015 correlated with ELA 2016, in bold), which is 0.630. This correlation is strong and positive and is higher than the two available discriminant validity coefficients (Science 2015 correlated with ELA 2016 (0.459), Science 2015 correlated with Math 2016 (0.440). This pattern of correlations represents strong validity evidence for PACE. This same pattern persists for all grade pairs except grades 7 and 8.

Table 2. Correlation Table Grade 4-5

	Science 2015	ELA 2015	ELA 2016
ELA 2015	*.555*		
ELA 2016	.459	**.630**	
Math 2016	.440	.603	.735

Table 3. Correlation Table Grade 5-6

	Math 2015	ELA 2015	ELA 2016
ELA 2015	*.635*		
ELA 2016	.520	**.619**	
Math 2016	**.625**	.590	*.648*

Table 4. Correlation Table Grade 6-7

	Math 2015	ELA 2015	ELA 2016
ELA 2015	*.470*		
ELA 2016	.482	**.586**	
Math 2016	**.558**	.513	*.531*

For grades 7 and 8 we see somewhat weaker convergent validity coefficients (0.483 for ELA, 0.488 for math). These are still strong validity coefficients, but are not as strong as for previous grade pairs. It is more telling that the discriminant validity coefficient for Math 2015 to ELA 2016 is higher than the convergent validity coefficients. This may indicate an update in task development, administration, or scoring that impacted the 2016 data and attenuated the correlations between like subjects. This analysis should be revisited next year to ensure that the correlations for this grade pair are similar to the other grade pairs and follows the expected pattern.

Table 5. Correlation Table Grade 7-8

	Math 2015	ELA 2015	Science 2016	ELA 2016
ELA 2015	*.517*			
Science 2016	.523	.448		
ELA 2016	.557	**.483**	*.574*	
Math 2016	**.488**	.477	*.590*	*.541*

Taken together, the correlation results provide strong evidence that PACE is functioning as intended. The correlations among the PACE subject areas within and across grades are similar to other statewide assessments. Correlations within year among the PACE subjects were quite high, especially for elementary grades.

Teacher Survey

Survey Response Rates

Each participating NH Tier I district provided a list of teacher names and email addresses that were to receive the survey. HumRRO standardized the file formats and resolved discrepant records (e.g., deleted duplicates, fixed transposed names).

Some districts included a position/title field along with the list of names and emails; others did not. Some of the positions/titles included for some of the districts appeared to include individuals who might not be teachers (e.g., guidance counselor, social worker). Consequently, with approval from PACE Leadership, after the initial branching item on the survey that identifies PACE teachers vs. "other" teachers, the following item was added to the Other Teachers Survey, "Are you currently teaching students in a particular grade level(s) or content area (e.g., kindergarten teacher, 6th grade social studies teacher, elementary music school teacher)?" If respondents selected "No," they were thanked for their time and directed to exit from the survey. All of the non-teachers were removed from the numerator and denominator in the calculation of the response rates.

Also, to be retained in the analyses, individuals had to have responded to items beyond the background items on the surveys. That is, respondents who logged on to the survey but did not complete any items or did not respond to any items beyond the background items, were removed. The number of cases reported in Table 6 (see columns titled "Num. Valid Reponses to PACE Teachers Survey" and "Num. Valid Responses to Other Teachers Survey) reflect the number of cases that met these criteria. Table 6 reveals that the overall response rate across all participating districts was 43%, which compares favorably to typical on-line survey response rates. The response rates varied considerably by district, with a high of 100% for Monroe (the smallest pilot district) to a low of 30% for Concord.[11]

Table 6. Overall Survey Response Rates by Participating PACE Tier 1 District

District	Num. Valid Emails [a]	Num. who Self-Identified as Not Teaching	Num. Valid Teacher Emails	Num. Valid Responses[b] to PACE Teachers Survey	Num. Valid Responses[b] to Other Teachers Survey	Final Overall Response Rate
Concord	387	22	365	67	41	29.6
Epping	89	10	79	23	33	70.9
Monroe	13	2	11	8	3	100.0
Pittsfield	62	10	52	16	19	67.3
Rochester	392	13	379	58	47	27.7
Sanborn	169	11	158	41	54	60.1
Seacoast	18	0	18	5	7	66.7
Souhegan	78	7	71	21	16	52.1
White Mountains	111	10	101	30	40	69.3
Sum	*1,319*	*85*	*1,234*	*269*	*260*	*42.9*

[a] Several email addresses were returned as undeliverable. In these instances, HumRRO contacted the PACE District Leads for corrections. In all but a few cases, valid emails were obtained for all the names provided in the master lists.
[b] To be retained for analyses, respondents had to respond to more than just the background items on the survey. Respondents who logged on to the survey but did not complete any items or did not respond to any items beyond the background items, were removed.

Background Characteristics of Respondents

The background characteristics of the respondents to the PACE Teachers Survey and the Other Teachers Survey are provided in Table 7. Table 8 provides additional background characteristics of the teachers who responded to the PACE Teachers Survey, including background information specifically about the teachers who administered NH PACE common performance tasks.

[11] It is worth noting that response rates may have been adversely impacted by the inclusion of "non-teachers" in the master email lists. To help mitigate this, we included a survey item that asked respondents to indicate if they taught students in a particular grade level(s) or content area so that we could remove them from the master list. However, these non-teachers had to actually log on to the survey and identify themselves as such in order to be removed from the response rate calculations. There may have been some non-teachers on the master list who did not log on to the survey to identify themselves as non-teachers. To the extent that this occurred, the response rates reported in Table 1 are downwardly biased.

Table 7. Background Characteristics of Respondents: PACE Teachers and Other Teachers

	PACE Teachers		Other Teachers	
Background Items	**n**	**%**	**n**	**%**
Grades Taught				
Kindergarten	-- [a]	--	54	20.8
Grade 1	--	--	64	24.6
Grade 2	--	--	65	25.0
Grade 3	43	16.0	43	16.5
Grade 4	50	18.6	41	15.8
Grade 5	52	19.3	43	16.5
Grade 6	37	13.8	41	15.8
Grade 7	35	13.0	44	16.9
Grade 8	26	9.7	50	19.2
Grade 9	64	23.8	68	26.2
Grade 10	78	29.0	73	28.1
Grade 11	58	21.6	90	34.6
Grade 12	54	20.1	85	32.7
Content Area				
ELA	135	50.2	98	37.7
Math	137	50.9	91	35.0
Science	90	33.5	76	29.2
Art	--	--	23	8.8
Music/Fine Arts	--	--	12	4.6
Social Studies/History	--	--	87	33.5
World Language	--	--	17	6.5
Special Education	--	--	12	4.6
Vocational Studies	--	--	10	3.8
Physical Ed/Health	--	--	13	5.0
Library	--	--	8	3.1
Computers/Technology	--	--	19	7.3
Other (including Drama)	--	--	24	9.2
Switch Schools or Districts				
No	241	89.6	219	84.2
Switched 2014–2015	6	2.2	9	3.5
Switched 2015–2016	13	4.8	13	5.0
Switched 2016–2017	14	5.2	23	8.8

[a] Dashes indicate that these response options did not appear on the PACE Teacher Survey. The PACE Teacher Survey was administered to those respondents who answered, "yes" to the initial branching item on the survey: "Did you/are you administering an NH PACE Common Performance Task(s) in mathematics, science, and/or ELA in grades 3-8 or high school anytime in 2015-16 or 2016-17?"

Table 8. Background Characteristics Specific to PACE Teachers Survey

Background Items	n	%
Length of time you have personally been a part of the PACE pilot program		
Since 2014–15 school year	133	49.4
Since 2015–16 school year	83	30.9
Since 2016–17 school year	53	19.7
Participated in cross-district collaborations		
Yes	182	67.7
No	87	32.3
PACE Content Lead		
Yes	29	15.9
No	153	84.1
Teacher Representative [a]		
Yes	116	63.7
No	66	36.3

[a] Defined in the survey as "participating in cross-district task development sessions and communicating progress on task development to other teachers in your district."

Overall Survey Results

The item-level frequency distributions for the PACE Teachers Survey and the Other Teachers Survey are provided in Appendices C and D, respectively. The frequency distributions are provided for all surveyed districts combined, and also disaggregated by district. However, to ensure anonymity of responses, the three smallest districts—Monroe, Seacoast, and Pittsfield—were collapsed into a "Small Districts Combined" category in the disaggregated reporting by district.

Most of the survey items were rated on a 5-point Likert scale, most often an agreement scale where 1 = Strongly disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; and 5 = Strongly agree. There were also a few items on the PACE Teachers Survey that were rated on a Usefulness scale (1 = Not useful; 2 = Slightly useful; 3 = Somewhat useful; 4 = Very useful; and 5 = Extremely useful) and one item on an Extent scale (1 = To no extent; 2 = To a slight extent; 3 = To some extent; 4 = To a great extent; and 5 = To a very great extent).

Overall, both the PACE teachers and the other teachers expressed favorable opinions about the PACE pilot program. There were no items on either the PACE Teachers Survey or the Other Teachers Survey that received a mean rating below 3.00 (i.e., the mid-point on the Likert scale) *or* that received more unfavorable ratings (i.e., 1s or 2s on the Likert scale) than favorable ratings (i.e., 4s and 5s).

The highest and lowest rated items, along with their descriptive statistics, are presented in Tables 4 and 5 for the PACE Teachers Survey and the Other Teachers Survey, respectively. It should be noted, as mentioned previously, that no items received a mean rating below 3.0. Consequently, the "lowest" rated items depicted in Tables 9 and 10 do not reflect unfavorably on the PACE pilot. Rather, these items simply received lower ratings relative to the other items

and, therefore, represent the greatest areas for improvement relative to other topics addressed by the surveys.

Table 9. Highest and Lowest Rated Items on the PACE Teachers Survey

Likert Scale Survey Items	Claim Addressed by Item	Mean Rating [a]	S.D.	% Selecting Strongly Disagree + Disagree	% Selecting Strongly Agree + Agree
Highest Rated (item mean > 4.00)					
Q6b. My school's administration provides me with the resources and supports that I need to effectively implement the NH PACE Common Performance Tasks.	1a	4.40	0.76	6.0	81.4
Q6d. The teachers at my school effectively collaborate with one another on topics relevant to the implementation of the PACE pilot.	1a	4.20	0.92	6.7	81.4
Q24a1. Implementing performance tasks has had a positive impact on instructional practice, such that instruction occurs at a higher depth of knowledge in my classroom.	3b	4.09	0.82	5.0	81.4
Q24b1. Implementing performance tasks has had a positive impact on student engagement while completing performance tasks in my classroom.	3c	4.06	0.92	6.0	81.4
Lowest Rated (item mean < 3.50)					
Q14d. Indicate the extent to which the Content Leads have provided useful guidance and support in answering questions about scaffolding.[b]	1a & 1b	3.45	1.11	15.3	46.1
Q25b. The scoring resources available on the LibGuide effectively explain how to score the student work on the NH PACE Common Performance Tasks.	4a	3.37	0.92	11.8	40.2
Q6e. The time and effort required by the PACE initiative are worth the benefits that I have experienced and/or seen.	1a	3.31	1.27	24.2	45.7
Q7. Select the statement that most closely reflects your perception of teachers' opinion of PACE at your school.[c]	1a	3.29	0.89	18.2	41.2
Q26. The NH PACE Common Performance Tasks are more accessible to a greater range of student learning needs (e.g., students with disabilities, English language learners) than traditional standardized tests.	2b	3.28	1.20	24.2	45.8
Q25a. The scoring rubrics for the NH PACE Common Performance Tasks are sufficiently clear and detailed to ensure that separate scorers scoring the same student work arrive at the same score.	4a	3.25	1.09	23.8	48.9

[a] For items with a "Don't know" and/or a "Not applicable" response option, these responses were recoded to "missing" prior to computing the means and standard deviations.
[b] This item (Q14d) rated on a 5-pt Usefulness scale where 1 = Not Useful; 2 = Slightly useful; 3 = Somewhat useful; 4 = Very useful and 5 = Extremely useful). All other items in Table 4 rated on an Agreement scale where 1 = Strongly disagree; 2 = Disagree; 3 = Neutral; 4 = Agree and 5 = Strongly agree.
[c] This item (Q7) rated on the following scale: 1= None have a favorable opinion of PACE; 2 = Few have a favorable opinion of PACE; 3 = Some have a favorable opinion of PACE; 4 = Most have a favorable opinion of PACE; 5 = All have a favorable opinion of PACE (item reverse coded from its appearance on the survey).

Table 10. Highest and Lowest Rated Items on the Other Teachers Survey

Likert Scale Survey Items	Claim Addressed by Item	Mean Rating[a]	S.D.	% Selecting Strongly Disagree + Disagree	% Selecting Strongly Agree + Agree
Highest Rated (item mean > 4.00)					
Q8. My school's administration (e.g., principal, assistant principal, curriculum director) is supportive of the PACE initiative.	1a	4.42	0.80	2.3	84.3
Q13a1. Implementing performance tasks has had a positive impact on instructional practice, such that instruction occurs at a higher depth of knowledge in my classroom.	3b	4.21	0.75	1.6	82.9
Q13b1.Implementing performance tasks has had a positive impact on student engagement while completing performance tasks in my classroom.	3c	4.17	0.83	2.3	81
Q13c1. Implementing performance tasks has had a positive impact on student engagement in learning overall in my classroom.	3c	4.13	0.81	2.4	78.6
Q13a2. Implementing performance tasks has had a positive impact on instructional practice, such that instruction occurs at a higher depth of knowledge in my school.	3b	4.09	0.79	2.4	70.8
Q13b2. Implementing performance tasks has had a positive impact on student engagement while completing performance tasks in my school.	3c	4.07	0.78	2.4	67.3
Q13c2. Implementing performance tasks has had a positive impact on student engagement in learning overall in my school.	3c	4.01	0.81	2.4	65.3
Lowest Rated (item mean < 3.50)					
Q5. Please rate your level of familiarity with the NH Performance Assessment of Competency Education (PACE) pilot program at your school[b]	1a	3.02	0.92	20.4	26.5

[a] For items with a "Don't know" and/or a "Not applicable" response option, these responses were recoded to "missing" prior to computing the means and standard deviations.
[b] This item was rated on a Familiarity scale where 1 = Unfamiliar; 2 = Somewhat Unfamiliar; 3 = Somewhat familiar; 4 = Very familiar; and 5 = Extremely familiar.

On both surveys, among the highest rated items were the support from the school's administration for PACE and the positive impact of PACE on instruction and student engagement. Findings from the PACE Teachers Survey indicate that potential areas for improvement may include increasing clarity on scaffolding, the LibGuide resources for scoring student work, and scoring rubrics. The accessibility of the NH PACE common performance tasks for students with a greater range of learning needs (e.g., students with disabilities, English language learners) was also identified as a concern by nearly a fourth of the respondents (Q26). Also, nearly a fourth of the teachers who responded to the PACE Teachers Survey disagreed or strongly disagreed that the time and effort required by the PACE initiative are worth the benefits they have seen and/or experienced (Q6e). This may partly help to explain why, when asked the degree to which teachers at their schools have a favorable opinion of PACE, more than a third

(35%) of the PACE teachers selected the middle rating of "Some teachers have a favorable opinion of PACE" (Q7).

The Other Teachers Survey, which contained considerably fewer survey items, did not have any items with a mean rating below 3.50, aside from the one item depicted in Table 5, which shows that other teachers identified themselves as *at least* somewhat familiar with the PACE initiative.

The final item on the PACE Teachers Survey asked respondents for any additional information they would like to share about PACE. The themes identified in teachers' open-ended comments were consistent with the lowest rated items on the PACE Teachers Survey. Of the 269 survey respondents, 92 (34.2%) provided an open-ended comment. Many of these comments included multiple issues; in total 264 discrete comments were identified. Of those, 44 were classified into a topic/theme that could be classified as a positive comment about PACE, 205 were coded into a topic/theme that could be classified as a negative comment about PACE, and 15 were classified as "miscellaneous/neutral." The open-ended comments were content analyzed by one HumRRO project team member and cross-checked by another HumRRO project team member. Discrepant codings were discussed and consensus was reached on all content codings.

Of the positive comments, the most frequent comments were general positive statements about PACE tasks (n = 13). The next most frequently mentioned themes were that PACE has a positive impact on instruction (n = 8) and that the collaboration is valued (n = 7).

The most frequently mentioned theme to emerge in the open-ended comments were concerns about the validity of the PACE scores (n = 33)[12]. Within this broad theme, specific concerns were mentioned regarding a need for more clarity in rubrics (n = 12), the accessibility of the tasks for special populations of students (n = 12), and confusion regarding scaffolding (n = 4); these same topics emerged amongst the lowest rated items on the survey. Additional context was provided by several teachers (n = 7) who explained they believe that the reading and writing demands required by the common tasks are too high. The next most frequently mentioned theme to emerge in the open-ended comments was a concern about too much time being spent on PACE requirements (n = 25). Teachers' comments ranged from general statements about the amount of work required to support PACE to specific comments about the amount of time spent developing and scoring tasks and the amount of time spent on the administrative aspects of PACE (e.g., organizing materials, collecting and scanning student work).

Additional detail on how the survey results inform the claims in the PACE Theory of Action is provided in the section of this report titled, "Summary of Findings by Evaluation Goal." In the next section, survey results disaggregated by key background variables are graphically depicted.

PACE Teachers Survey Results Disaggregated by Background Variables

During the course of this formative evaluation—through conversations with the PACE Leadership team, attendance at PACE meetings and task development sessions, and through interviews and focus groups with stakeholders—the potential for particular background variables to impact perceptions of the PACE pilot program were identified. While not practically feasible to conduct disaggregated survey analyses on all these background variables, we did investigate

[12] Most teachers who entered a response to the open-ended item provided comments on more than one topic. The average number of topics mentioned by teachers was 2.87.

some key background variables: time involved in the PACE pilot program (with special emphasis on comparing White Mountains with other districts), whether or not teachers had switched schools or districts since 2014–15, participation in cross-district collaboration, grade band taught (elementary, middle, and high school), and content area (ELA, mathematics and science).

Survey items that differed statistically ($p < .05$) or substantively across these background characteristics are graphically presented below. A substantive difference was defined as a mean effect size difference greater than or equal to $d = 0.40$; that is, an effect size that is greater than what is typically considered a small effect (Cohen, 1988). (A common rule of thumb for magnitudes of effect sizes is $d = 0.20$, small; $d = 0.50$, medium; $d = 0.80$, large; Cohen, 1988.) Many of the survey items included a "Don't know" and/or a "Not applicable" response option. Consequently, the statistical significance test was useful for flagging any items for which a disproportionate number of respondents from one category selected those options. These responses were omitted from mean rating calculations. The effect sizes provide a useful indication of the magnitude of the mean differences between the categories of respondents.

The items flagged for statistical or substantive differences are presented in figures by the claim they address in the PACE Theory of Action.[13] Items addressing different claims and/or items rated on different scales are presented in separate figures. Flagged items that address the same claim and rated on the same scale are depicted in the same figure. When multiple items are included in the same figure, the items are sorted by the items with the highest percentage of favorable ratings (e.g., Strongly agree + Agree) according to the category that received the overall highest ratings.

All of these figures use a common color scheme. Shades of green indicate ratings on the positive end of the response scale (Agree/Strongly Agree, Very Useful/Extremely Useful, To a Great Extent/To a Very Great Extent, etc.). Shades of orange/red indicate ratings on the negative end of the response scale (Strongly Disagree/Disagree, Not Useful/Slightly Useful, To No Extent/To a Slight Extent, etc.). In both cases, the more saturated color indicates a stronger response. Responses at the mid-point of the scale (Neutral, To Some Extent, Somewhat Useful, etc.) are in yellow. Responses of Don't Know or Not Applicable are in grey.

White Mountains Compared to Other Districts

The White Mountains School District represents a unique case in that it was the sole district to join the Tier 1 districts in 2016–17. Our teacher surveys were administered in November–December 2016, so many of the teachers might not yet have any substantive experience with PACE. Given their newness to PACE it is likely that their experiences and perceptions of the PACE program may have differed from those districts that had more familiarity. For this reason, we investigated whether there were statistical or substantive differences in responses for teachers from the White Mountains School District as compared to PACE teachers from the other Tier 1 districts. The items flagged for such differences are presented in Figures 4–10.

A trend that emerges across the figures is that PACE teachers in other districts tended to have a more favorable impression of PACE than PACE teachers in the White Mountains district on

[13] When the items for the survey were developed they were mapped to the claims from the theory of action that they most closely addressed. There is likely some overlap between items and claims such that some items may address multiple claims. Moreover, the connection between claims and items is more direct for some items than for other items.

these items. Two items in particular stand out for their large differences. Figure 5 (q6e) shows that PACE teachers in other districts were much more likely than PACE teachers in the White Mountains district to indicate that the teachers at their school had a favorable opinion of PACE (*d* = 0.75), and Figure 9 (q20a) shows that PACE teachers in other districts were much more likely to indicate that they received effective training and preparation to administer NH PACE common tasks (*d* = 0.96). Another trend that emerges across several of the figures is that the White Mountains PACE teachers were more likely to respond "Don't know" to several of the items. This is not surprising given that this is their first year of involvement in the PACE Tier 1 pilot program. Interestingly, there was one item that was flagged for which White Mountains PACE teachers had a higher rating than PACE teachers in the other Tier 1 districts (see Figure 9); White Mountains PACE teachers endorsed the item "competency-based education is integrated into my instruction" (q21) to a greater extent than other teachers (*d* = -0.40).

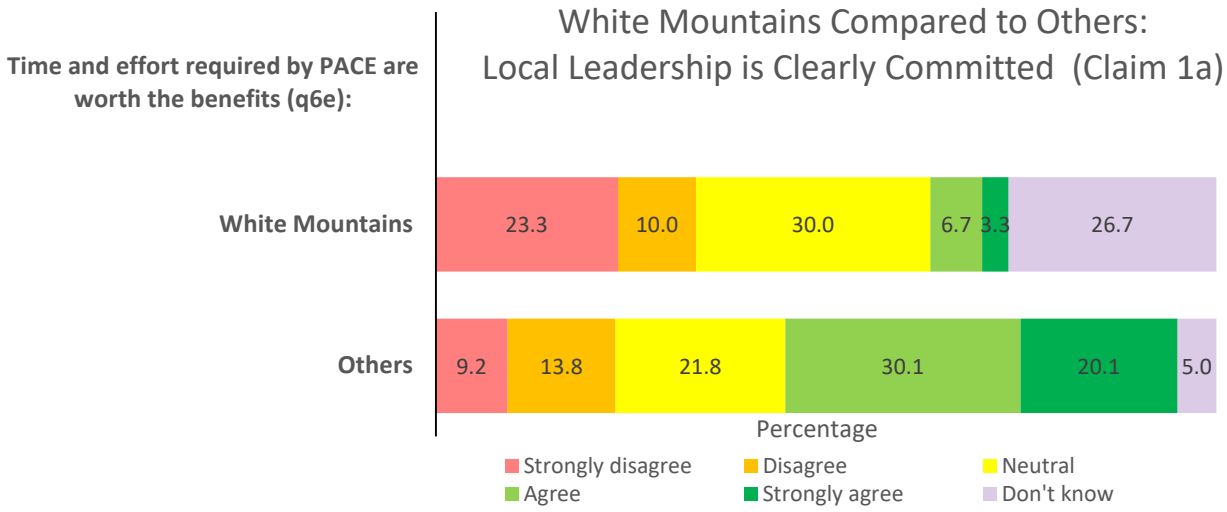


Figure 4. White Mountains v. other districts: Local leadership is clearly committed (q6).

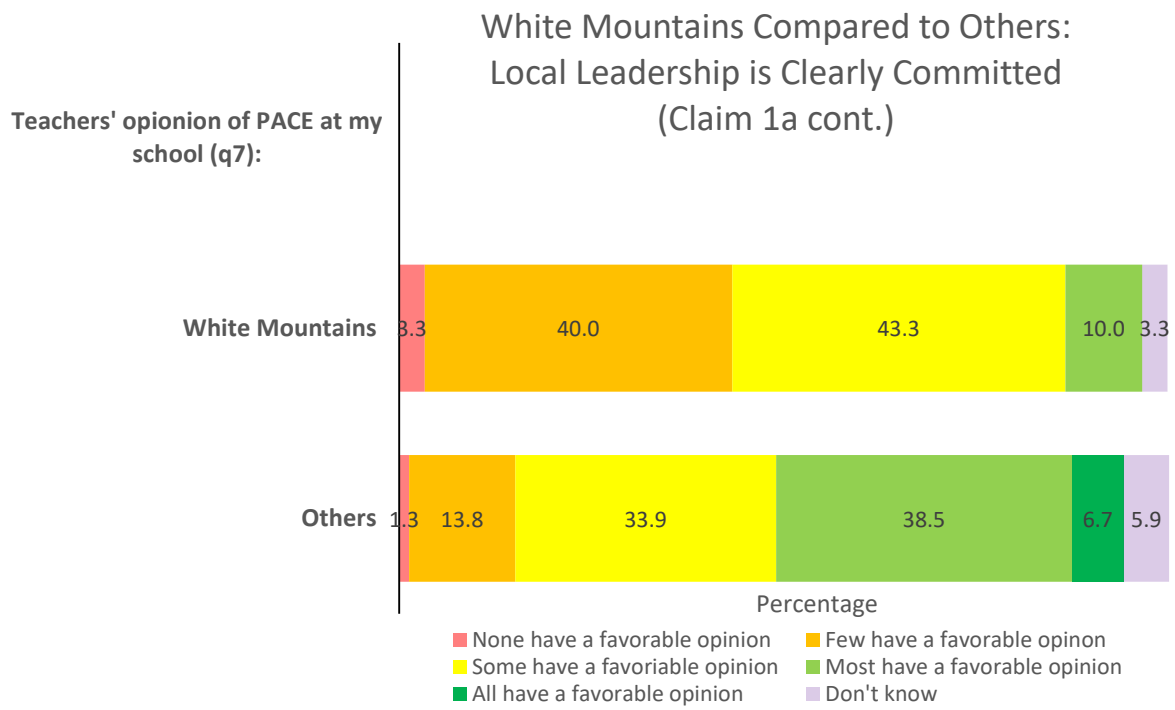


Figure 5. White Mountains v. other Tier 1 districts: Local leadership is clearly committed cont. (q7).

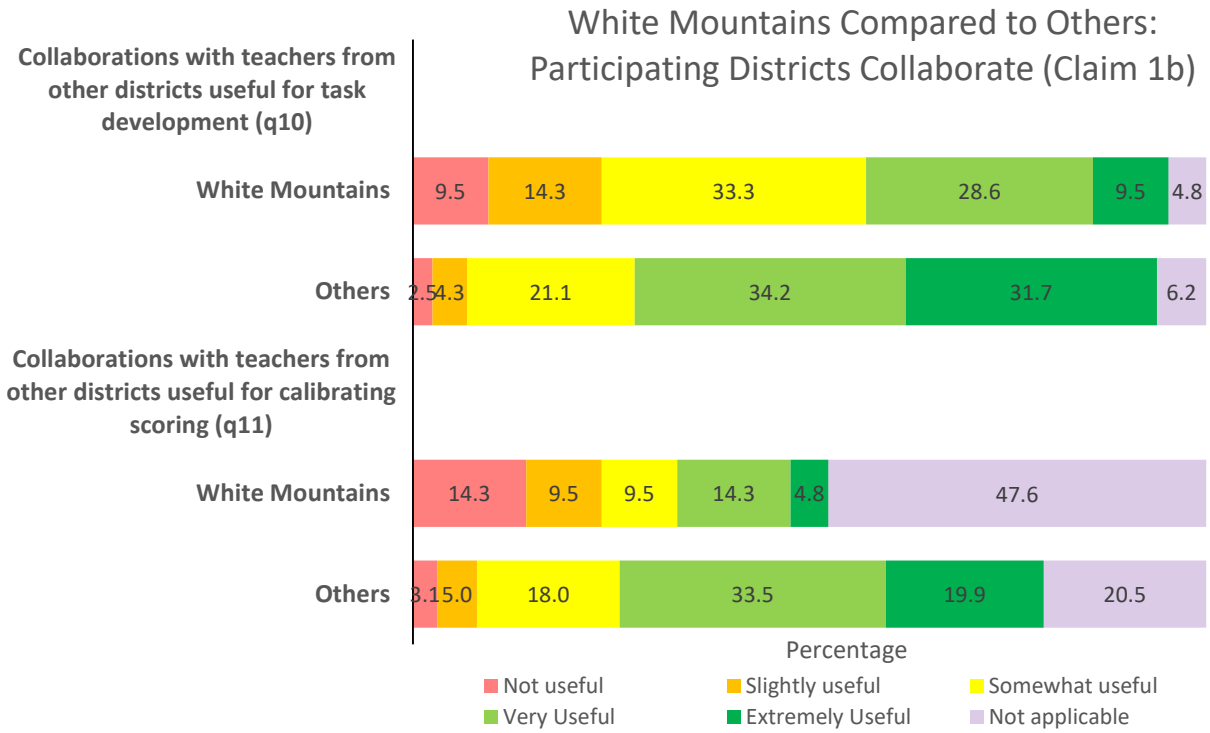


Figure 6. White Mountains v. other Tier 1 districts: Participating districts collaborate (q10 & q11).

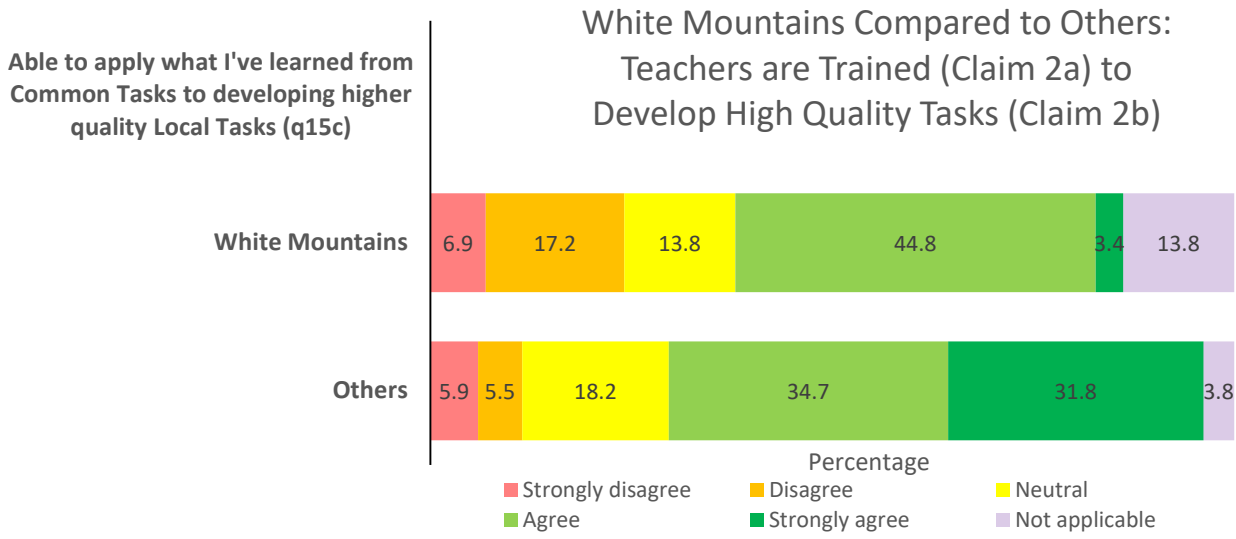


Figure 7. White Mountains v. other Tier 1 districts: Teachers trained to develop high quality tasks (q15c).

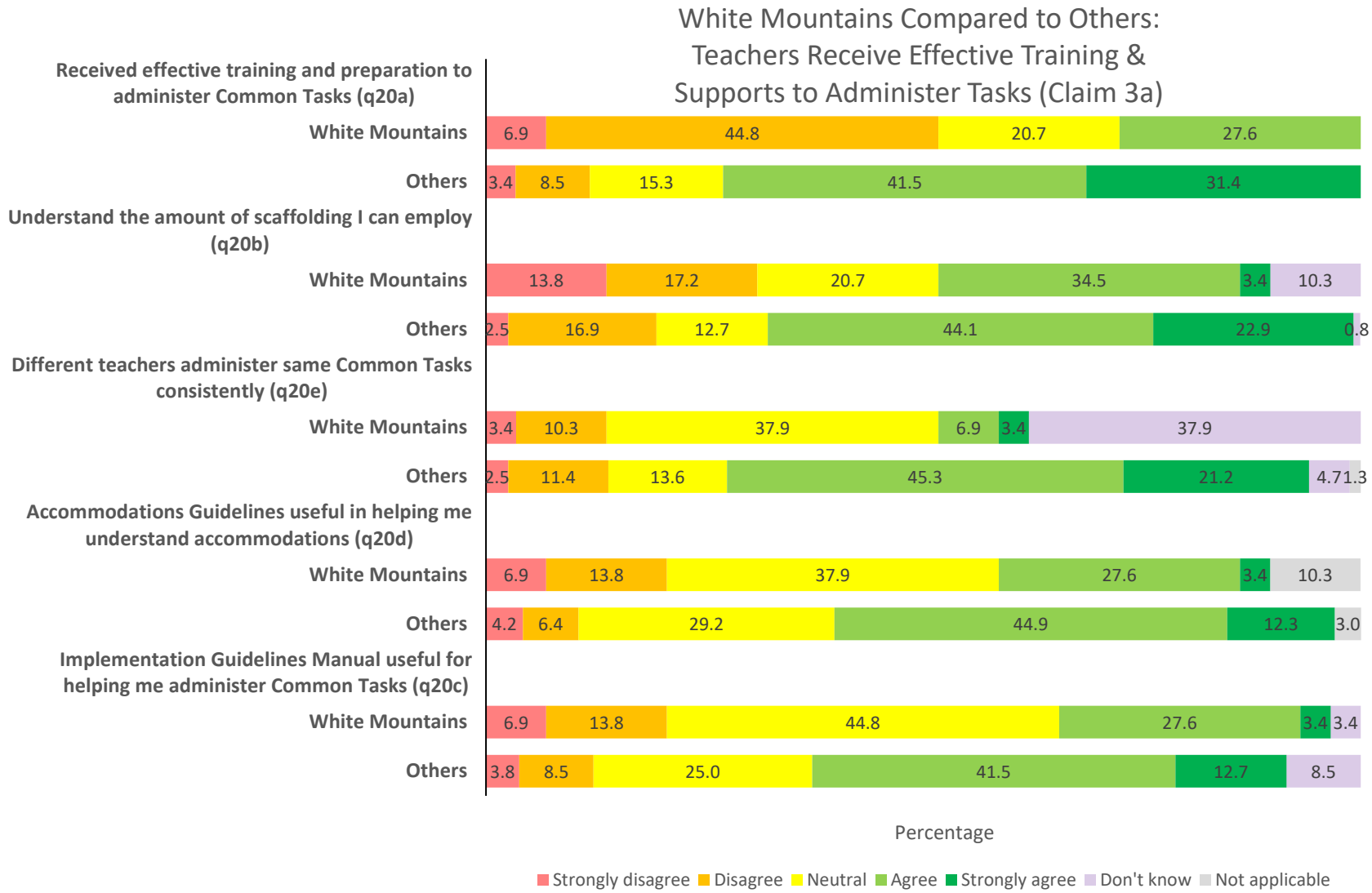


Figure 8. White Mountains v. other Tier 1 districts: Teachers receive effective training and supports to administer tasks (q20)

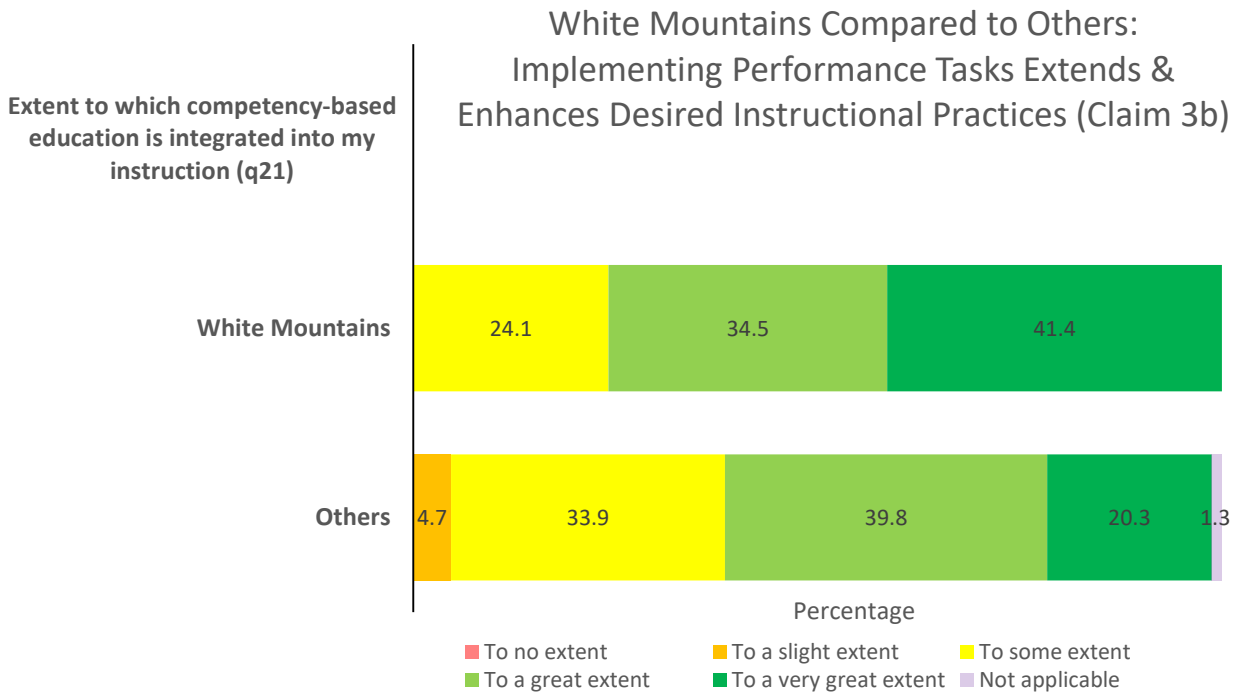


Figure 9. White Mountains v. other Tier 1 districts: Implementing tasks extends and enhances instructional practices (q21).

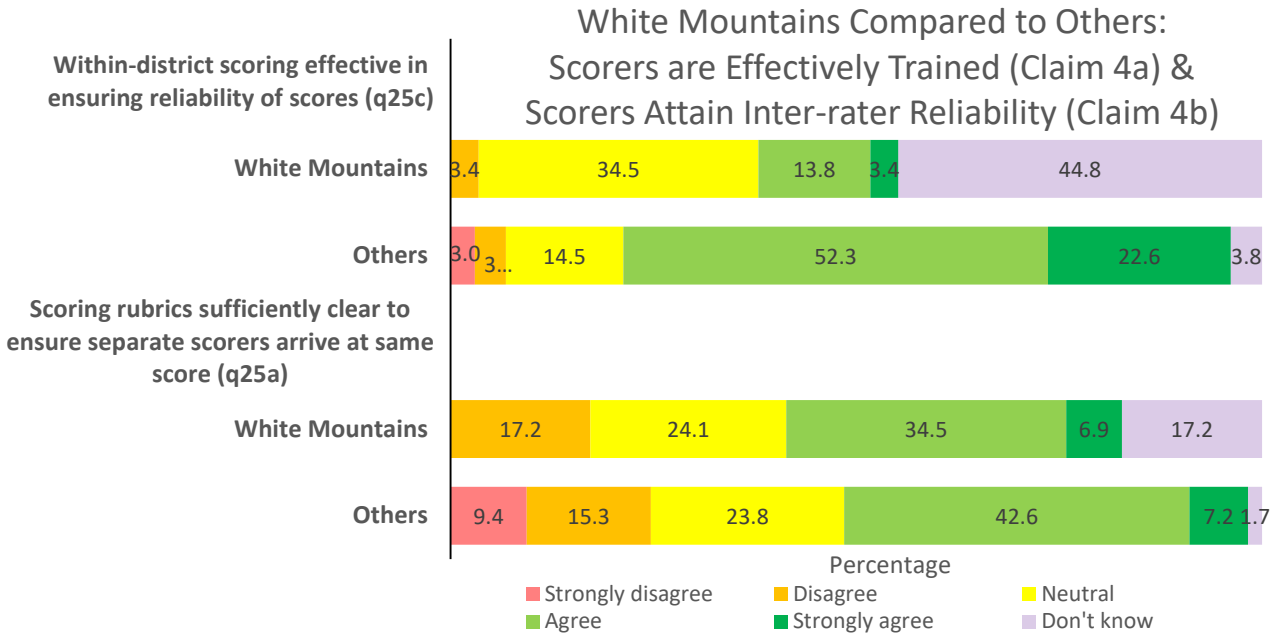


Figure 10. White Mountains v. other Tier 1 districts: Scorers are effectively trained & attain Inter-rater reliability (q25).

Switched Schools or Districts since 2014–15

Switching schools or districts may have an influence on one's perception of PACE for many possible reasons. For example, a new teacher in a PACE Tier 1 school might have less training on, and familiarity with, PACE principles and activities, and thus be more unsure of PACE. Alternatively, a teacher might transfer to a PACE school because it is a PACE school, and be highly favorable.

While the number of PACE teachers in the participating Tier I districts who indicated that they switched schools or districts was relatively small (n = 33), we investigated whether this group exhibited a notably different pattern of responses. A few items were for flagged for differences using the criteria noted above. Those items are presented in Figures 9—13. Across all these figures, there is a consistent trend for PACE teachers in the participating Tier I districts who switched schools or districts since 2014–15 to rate the items less favorably than those who did not switch schools or districts. The largest difference occurred for the item, "I have received adequate training and preparation to effectively administer NH PACE Common Tasks" (see Figure 14), such that teachers who were in the same school or district since 2014–15 were more likely to favorably endorse this item (*d* = 0.63).

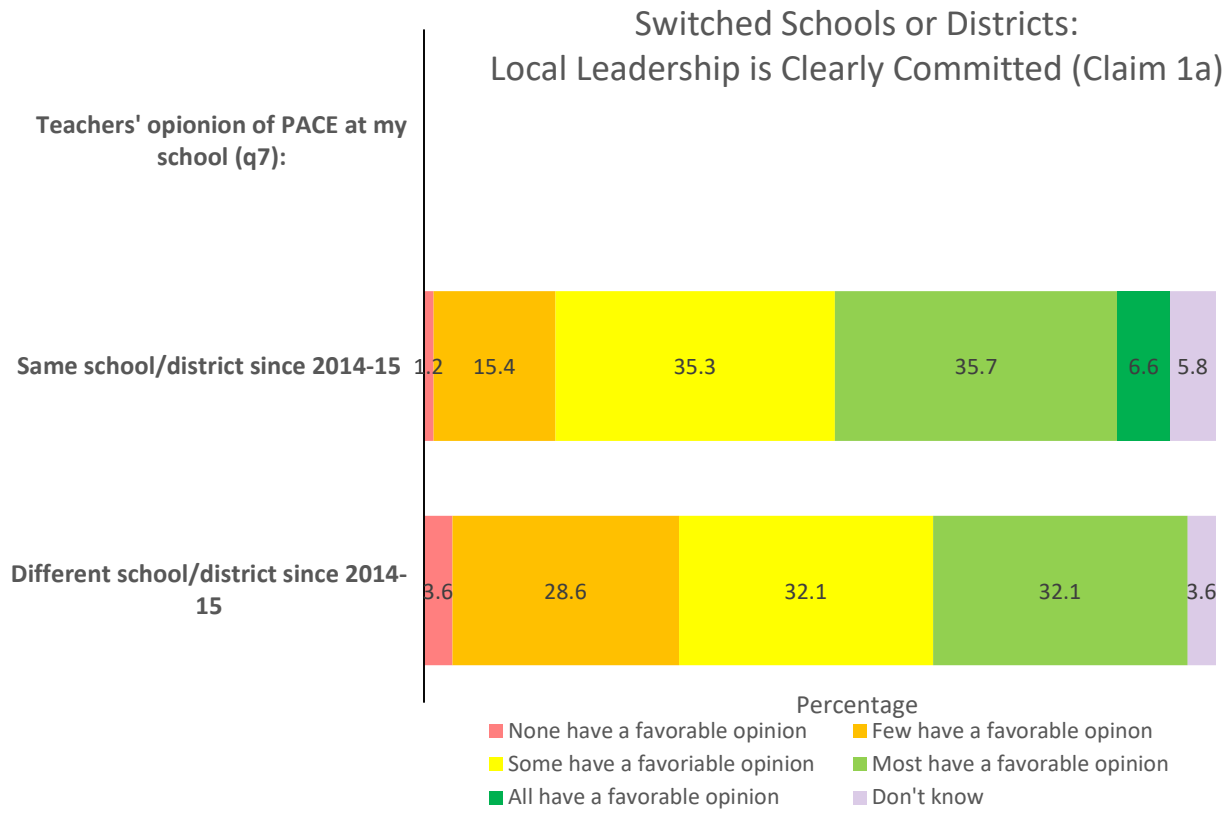


Figure 11. Switched schools/districts v. same school/district: Local leadership is clearly committed (q7).

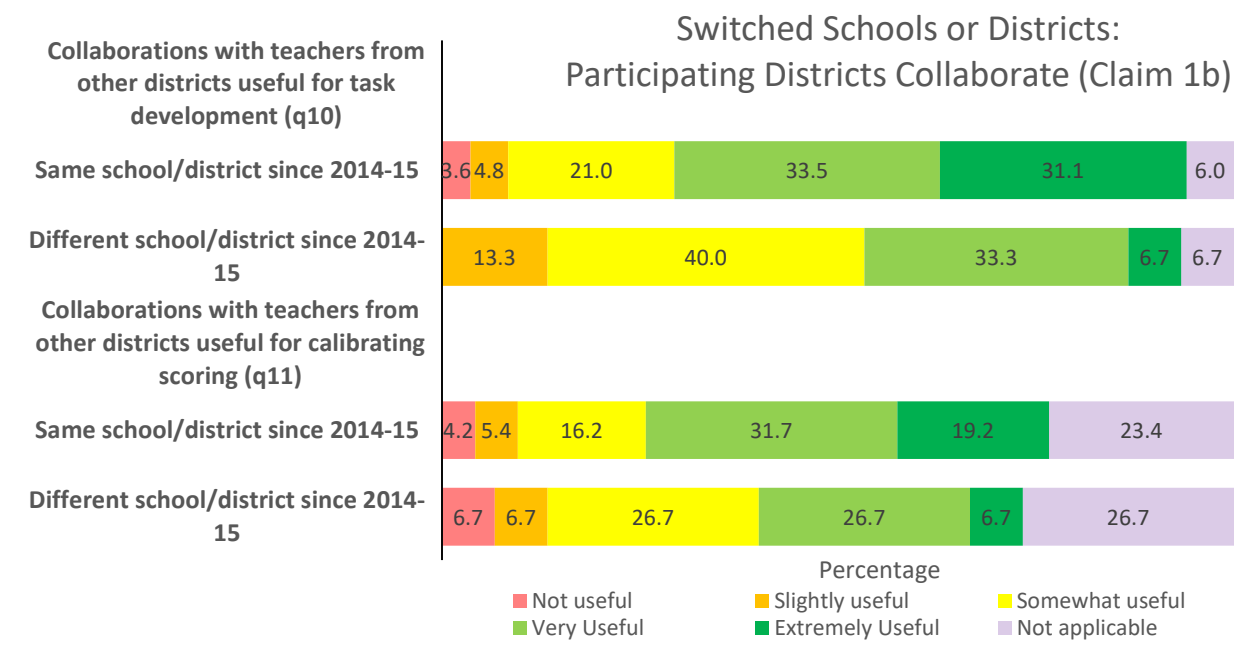


Figure 12. Switched schools/districts v. same school/district: Participating districts collaborate (q10 & q11).

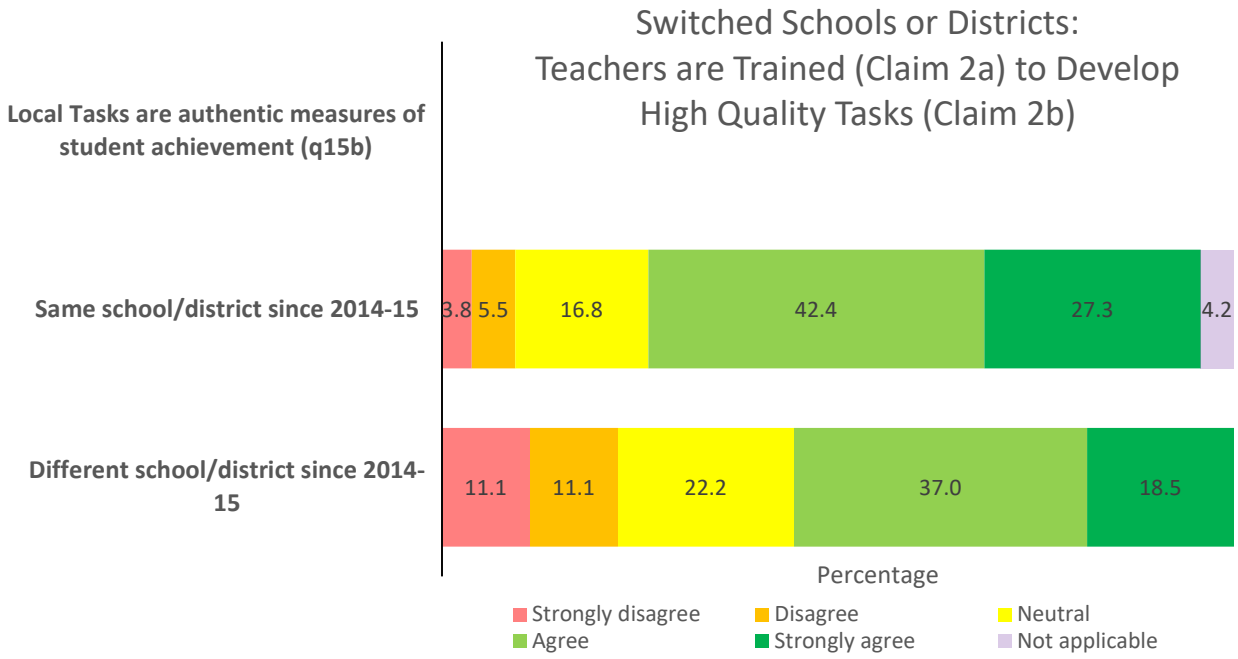


Figure 13. Switched schools/districts v. same school/district: Teachers are trained to develop high quality tasks (q15b).

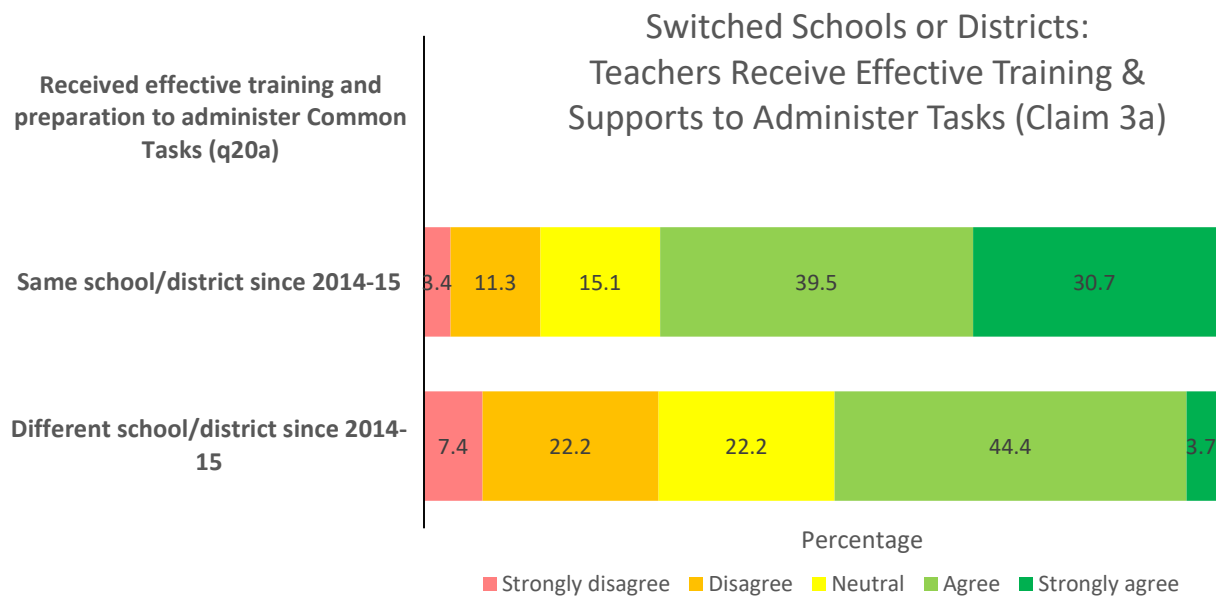


Figure 14. Switched schools/districts v. same school/district: Received effective training and preparation to administer tasks (q20a).

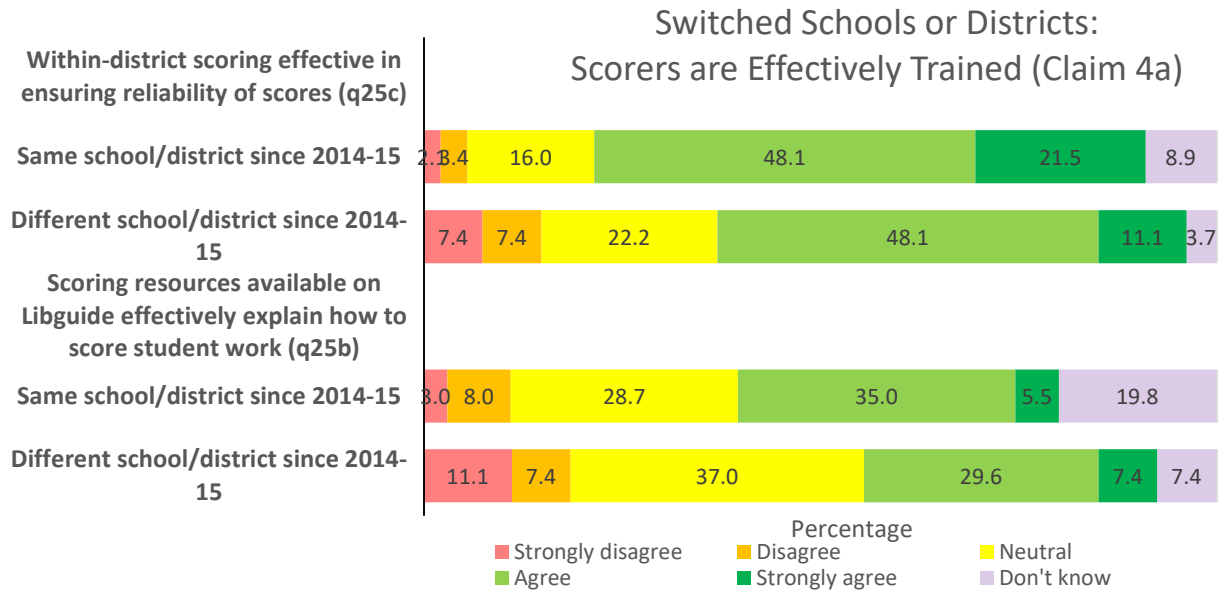


Figure 15. Switched schools/districts v. same school/district: Scorers are effectively trained (q25).

Cross-District Collaboration

An item on the PACE Teachers Survey asked respondents to indicate if they participated in cross-district collaborations on NH PACE common performance tasks (e.g., task development sessions and/or calibration sessions). Differences were investigated in responses for those who said "yes" (n = 182) and those who said "no" (n = 87). Several items were flagged for differences using the criteria previously noted.[14] Those items are presented in Figures 14—20. Across the flagged items, while both groups provided mostly favorable ratings, there was a consistent trend for those who participated in cross-district collaboration to have a more favorable impression of PACE. The most notable difference was for the item, "I have been able to apply what I've learned from the NH PACE Common Performance Tasks to developing higher quality local performance tasks" (see Figure 18, q15c). Those who participated in cross-district collaboration were considerably more likely to positively endorse this item (*d* = 0.73).

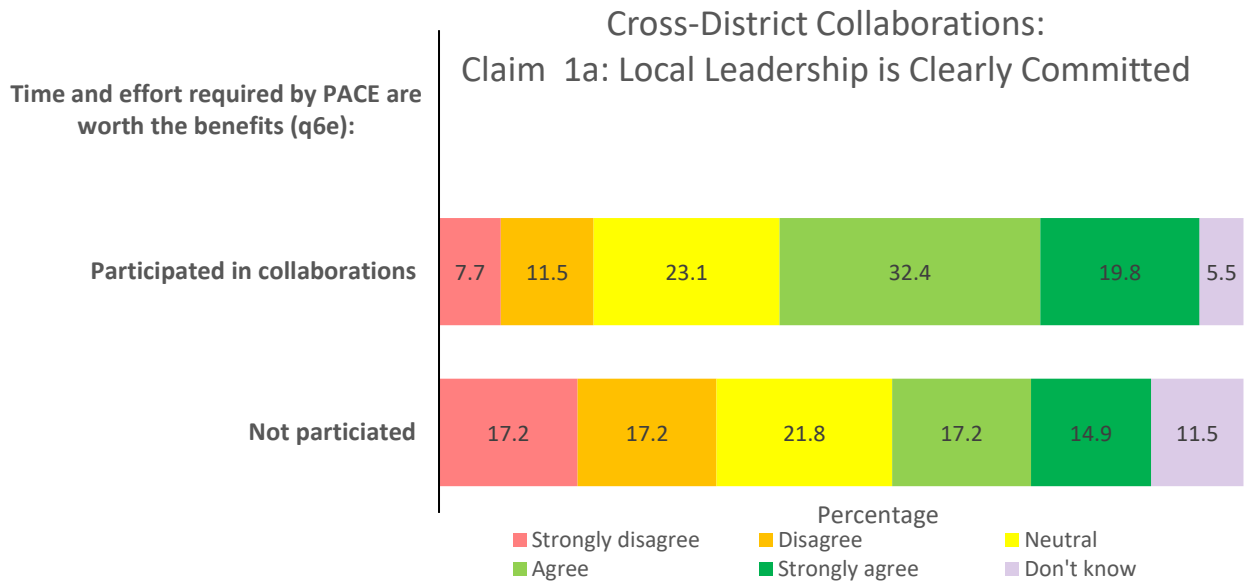


Figure 16. Participated in cross-district collaboration v. not participated: Local leadership is clearly committed (q6e).

[14] Items 9—14 on the PACE Teachers Survey were only presented to teachers who indicated that they had participated in cross-district collaborations (i.e., branching logic). Consequently, no comparisons were possible between these two respondent groups on this set of items.

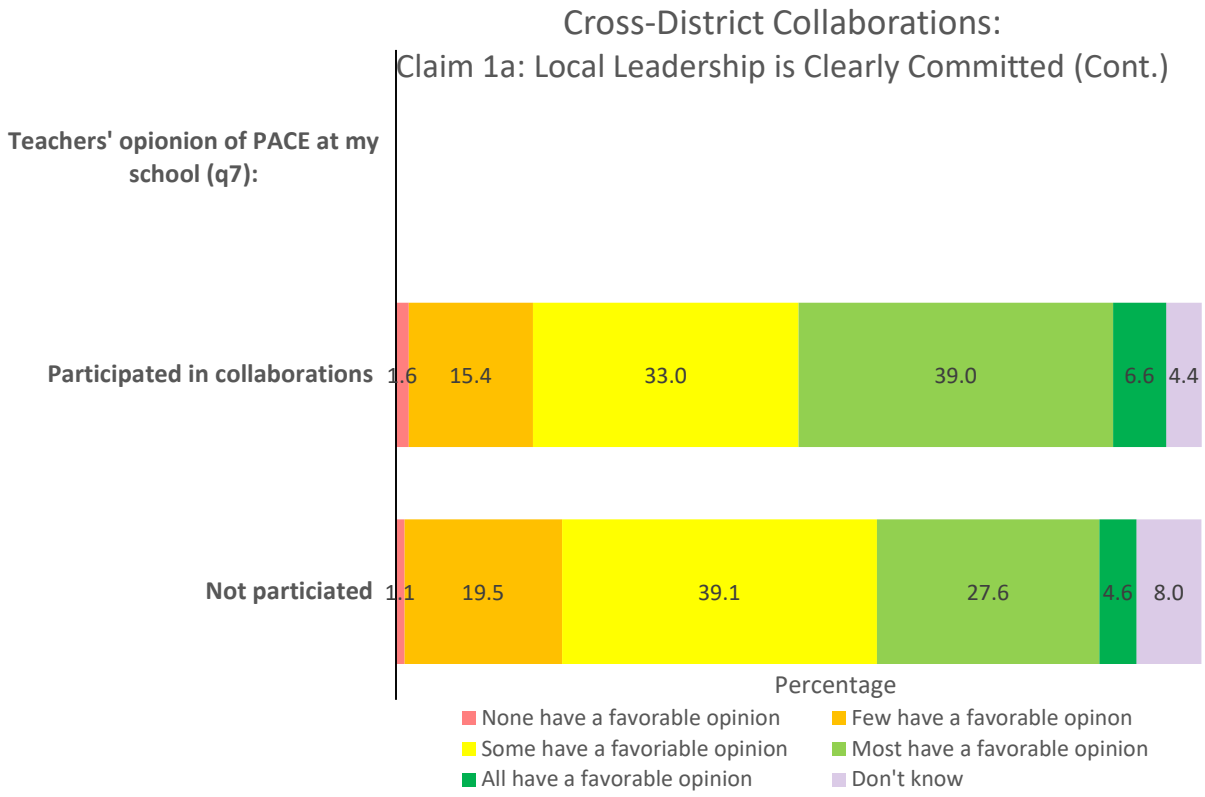


Figure 17. Participated in cross-district collaboration v. not participated: Local leadership is clearly committed cont. (q7).

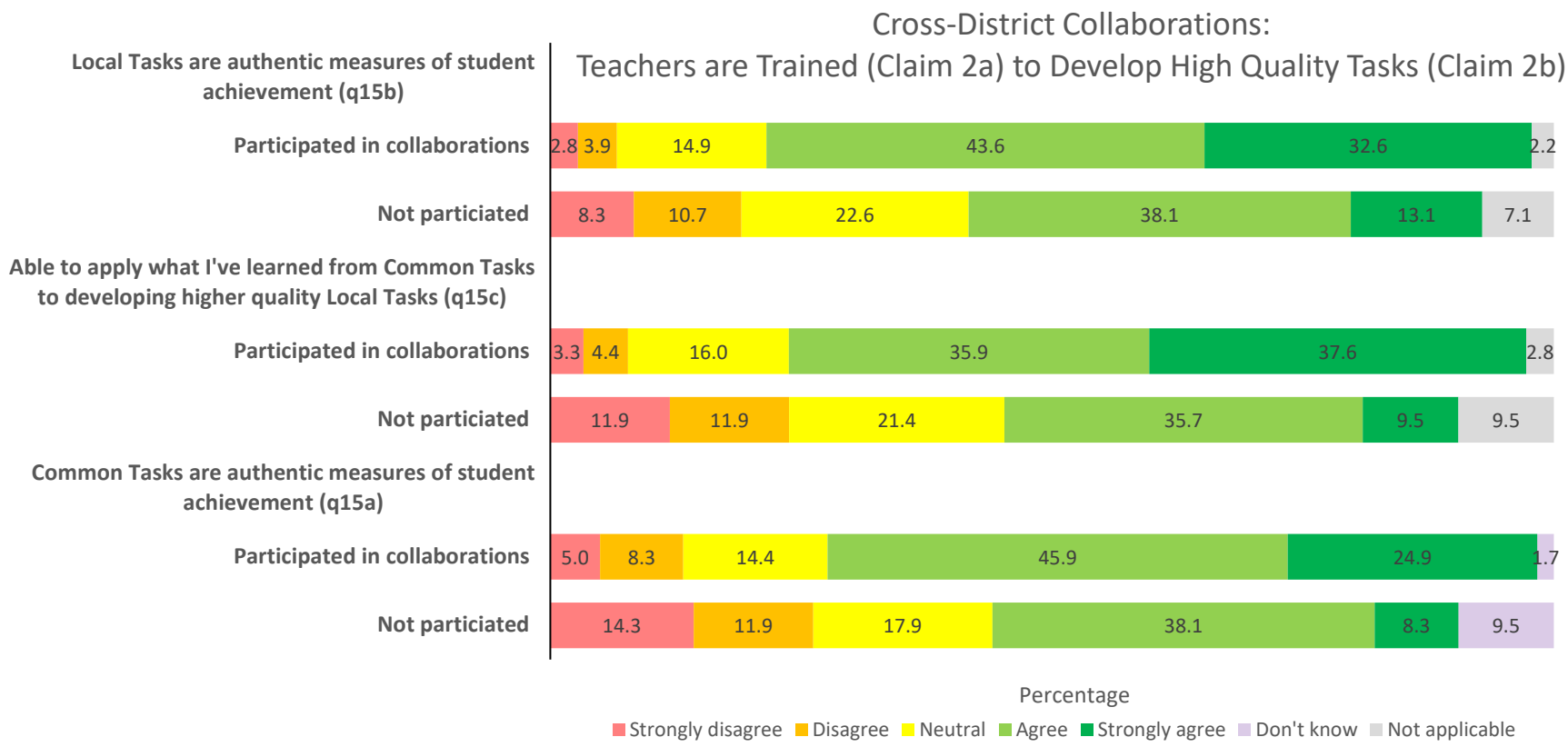


Figure 18. Participated in cross-district collaboration v. not participated: Teachers are trained to develop high quality tasks (q15).

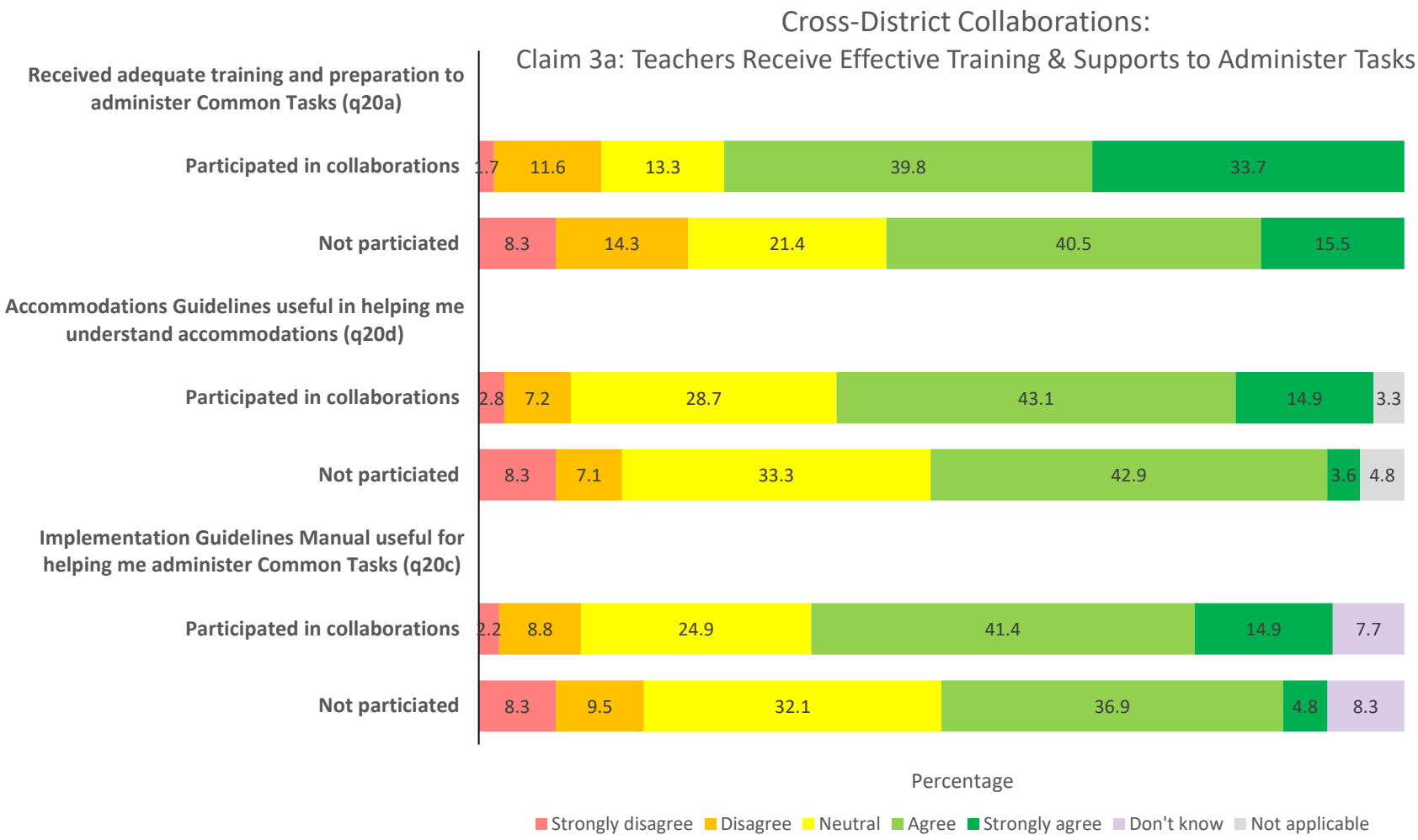


Figure 19. Participated in cross-district collaboration v. not participated: Teachers receive effective training and supports to administer tasks (q20).

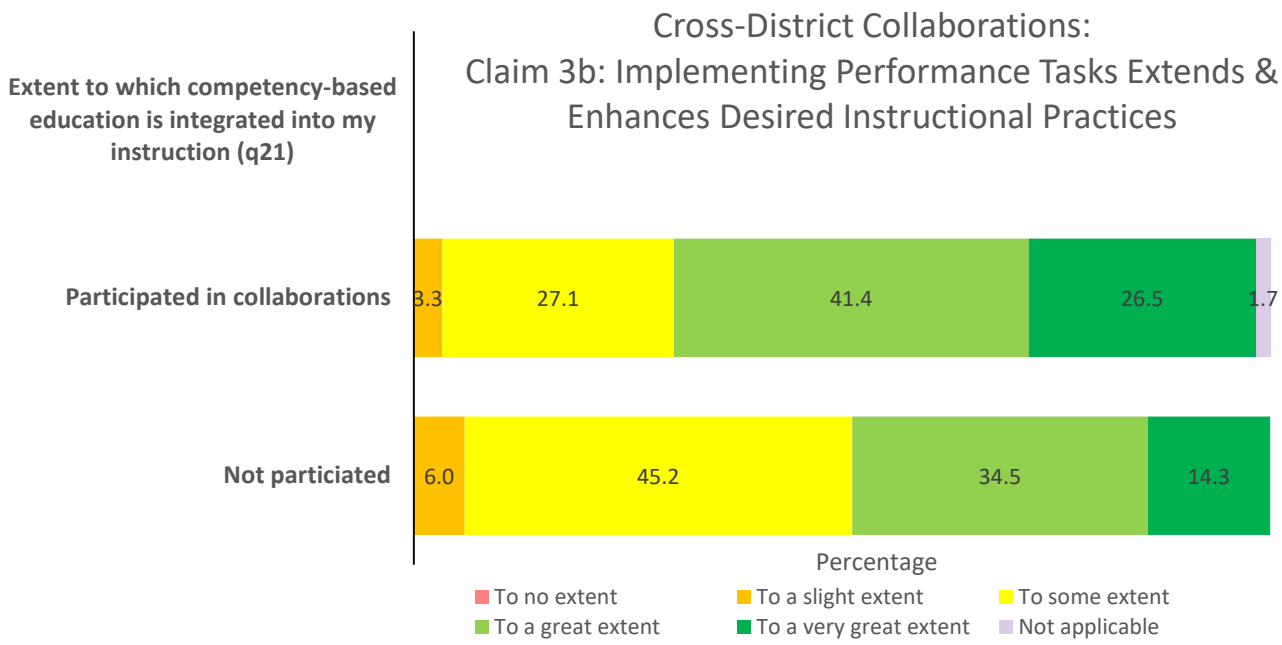


Figure 20. Participated in cross-district collaboration v. not participated: Implementing performance tasks extends and enhances instructional practices (q21).

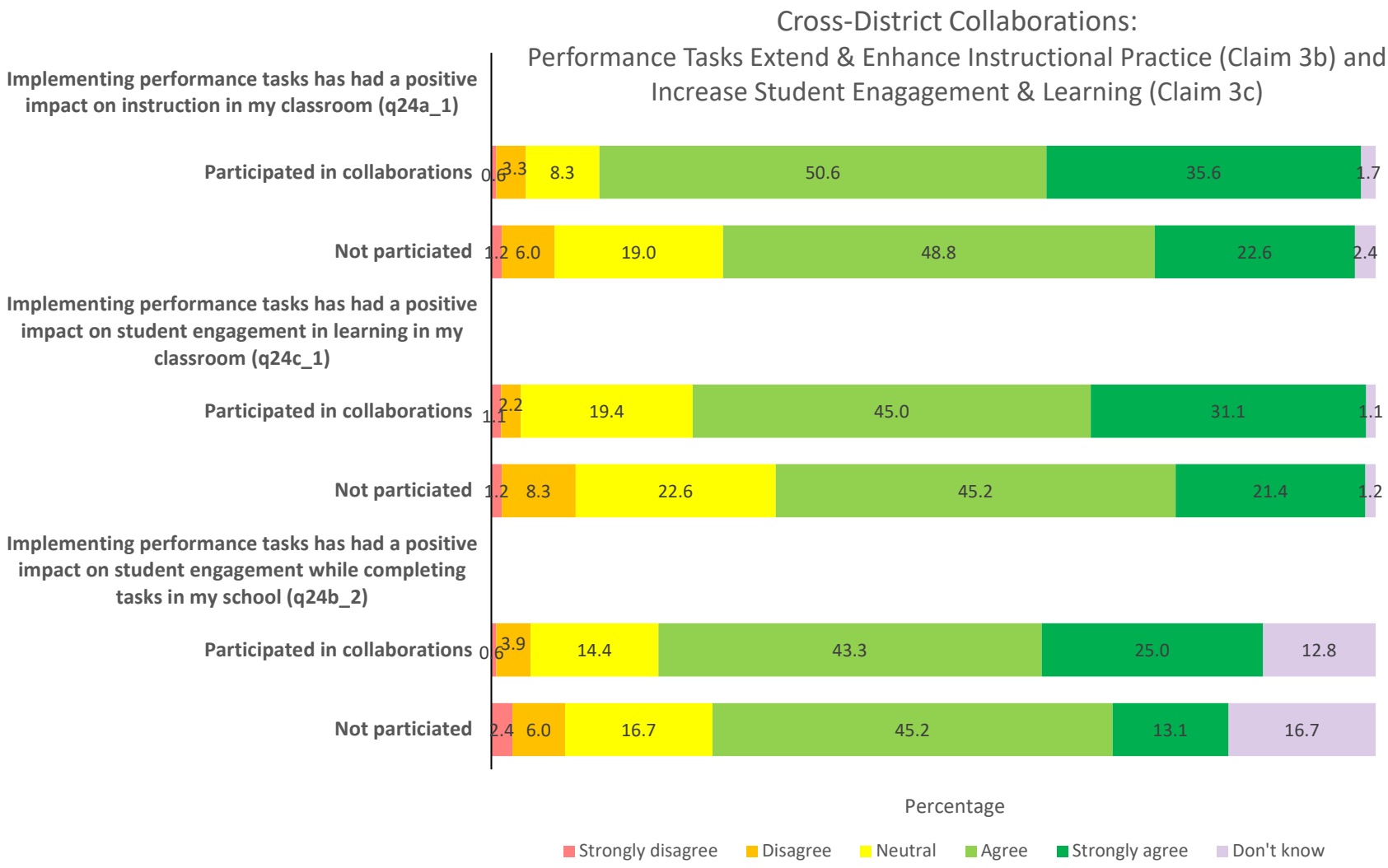


Figure 21. Participated in cross-district collaboration v. not participated: Tasks extend & enhance instructional practice and increase student engagement & learning (q24).

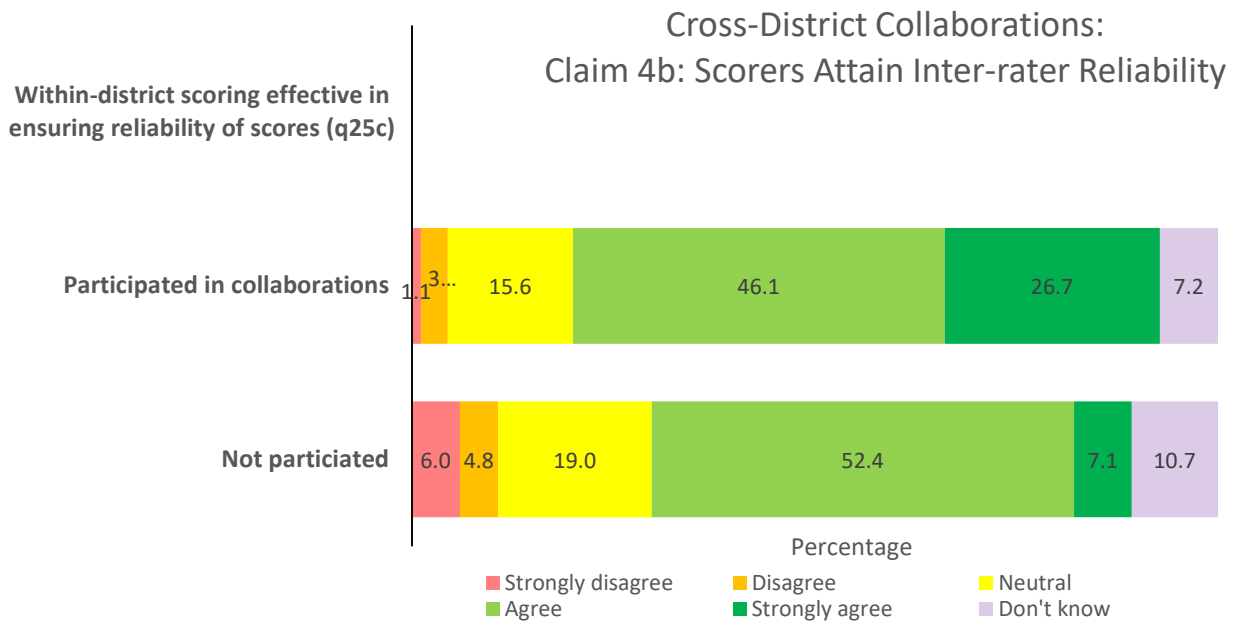


Figure 22. Participated in cross-district collaboration v. not participated: Scorers are effectively trained (q25c).

Grade Bands

Next, we identified respondents who administered NH PACE common tasks in elementary school (n = 112), middle school (n = 70), and high school (n = 100).[15] Differences in response patterns across these grade levels were investigated. The subset of items that were flagged for notable differences are presented in Figures 21–27.

Ratings for all three groups were generally quite positive. There was a consistent trend across most of the survey items such that high school PACE teachers in the participating Tier I districts tended to rate the items most favorably and middle school teachers tended to rate the items least favorably. The exception was for the set of items about the impact of PACE on instruction, student engagement, and student learning. Elementary school teachers tended to rate these items most favorably, followed by high school teachers and, finally, middle school teachers (see Figure 29). However, the magnitudes of the differences between elementary and high school teachers on these items were small (d = 0.14 to d = 0.30).

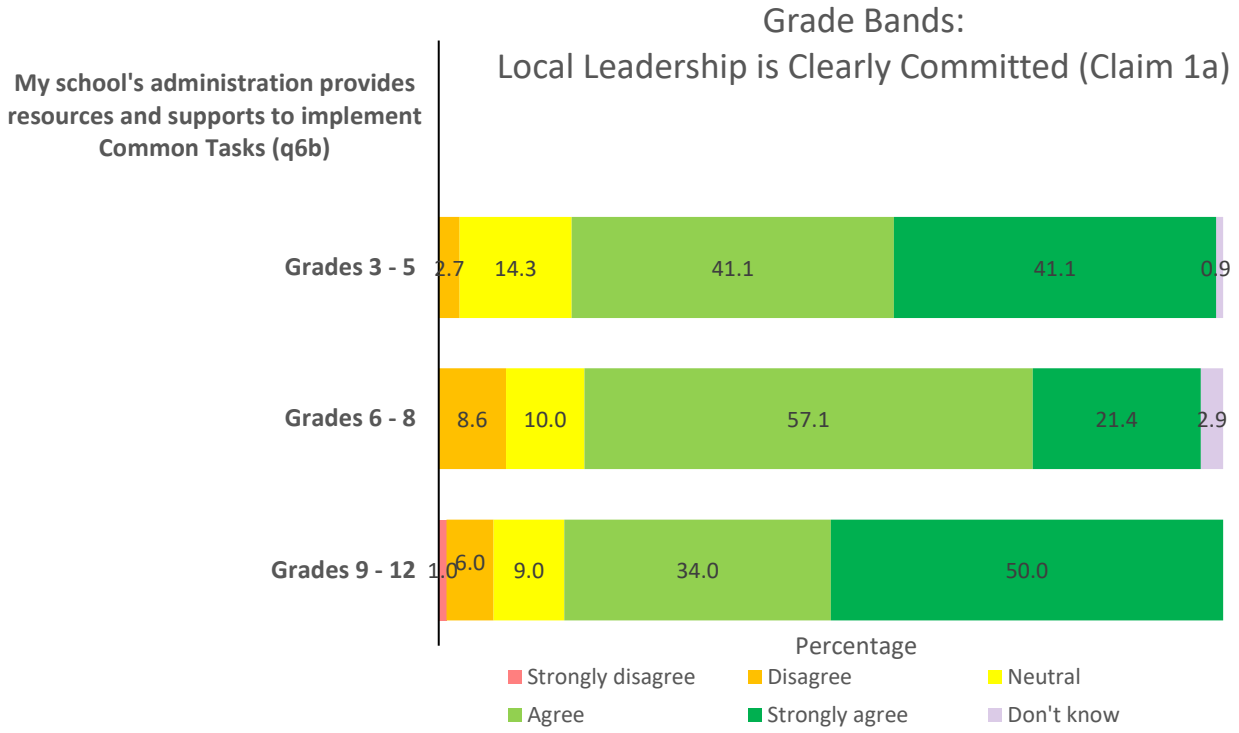


Figure 23. Grade bands comparison: Local leadership is clearly committed (q6b).

[15] Some teachers teach across grade bands (e.g., elementary school grades and middle school grades), which is why the number of teachers teaching elementary, middle, and high school is more than 269 (i.e., the total number of respondents to the PACE Teacher Survey). These individuals are represented multiple times in this section of the report.

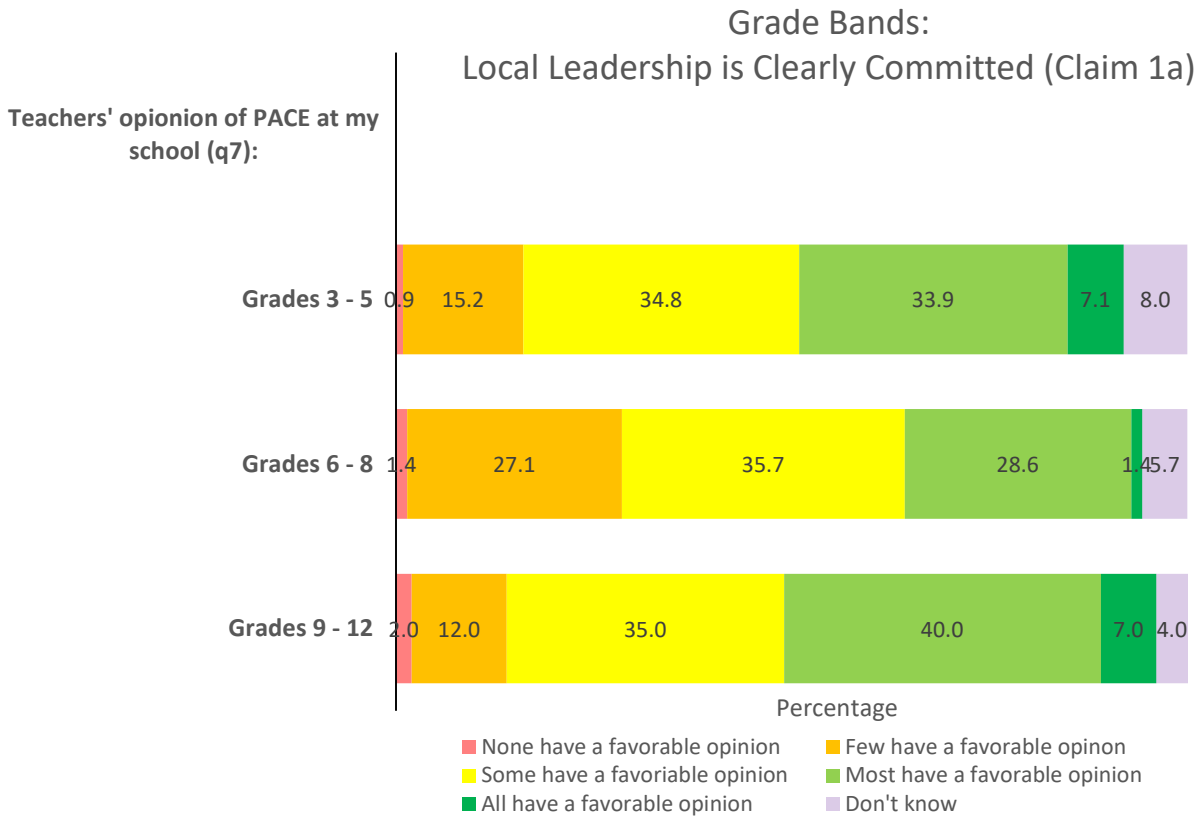


Figure 24. Grade bands comparison: Local leadership is clearly committed cont. (q7).

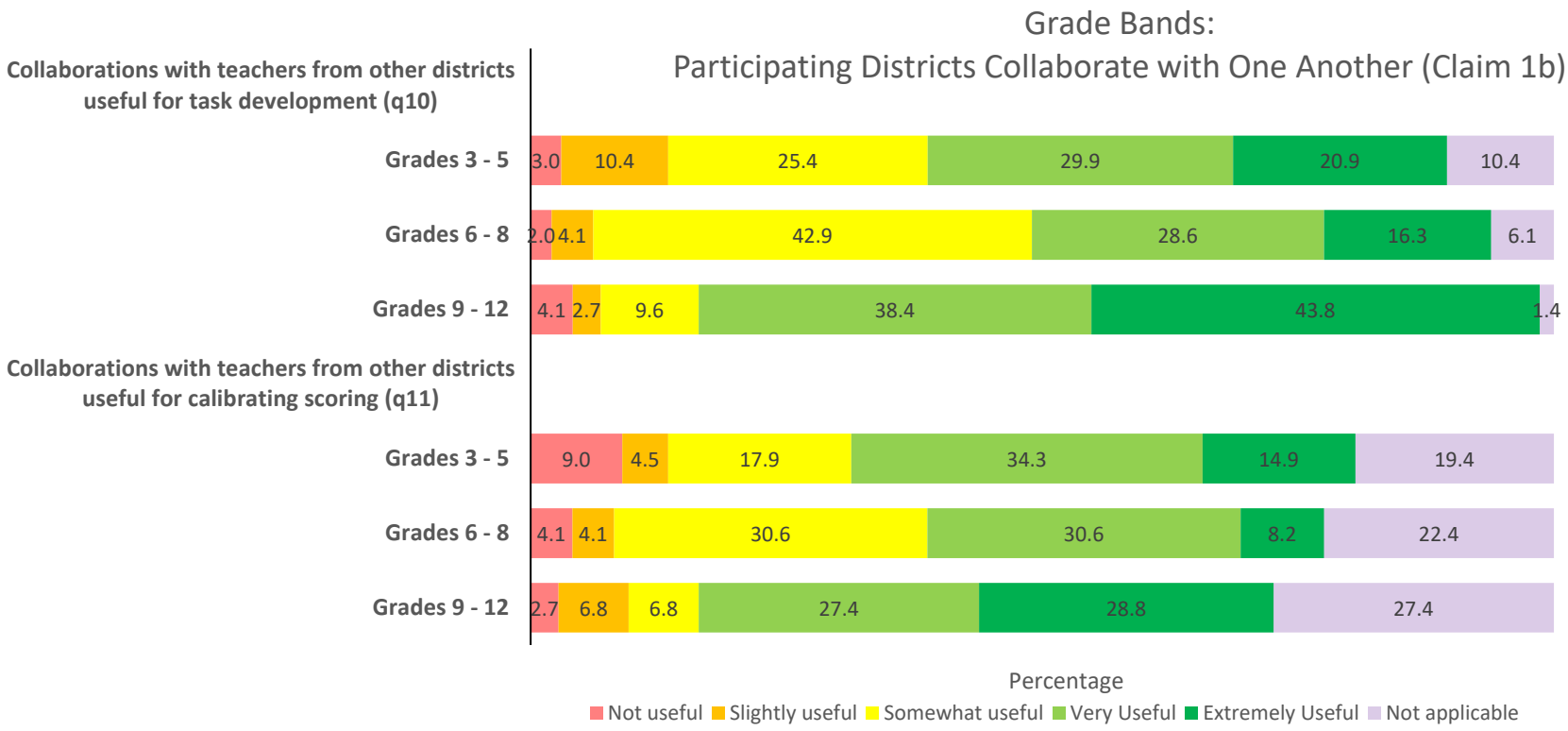


Figure 25. Grade bands comparison: Participating districts collaborate with one another (q10 & q11).

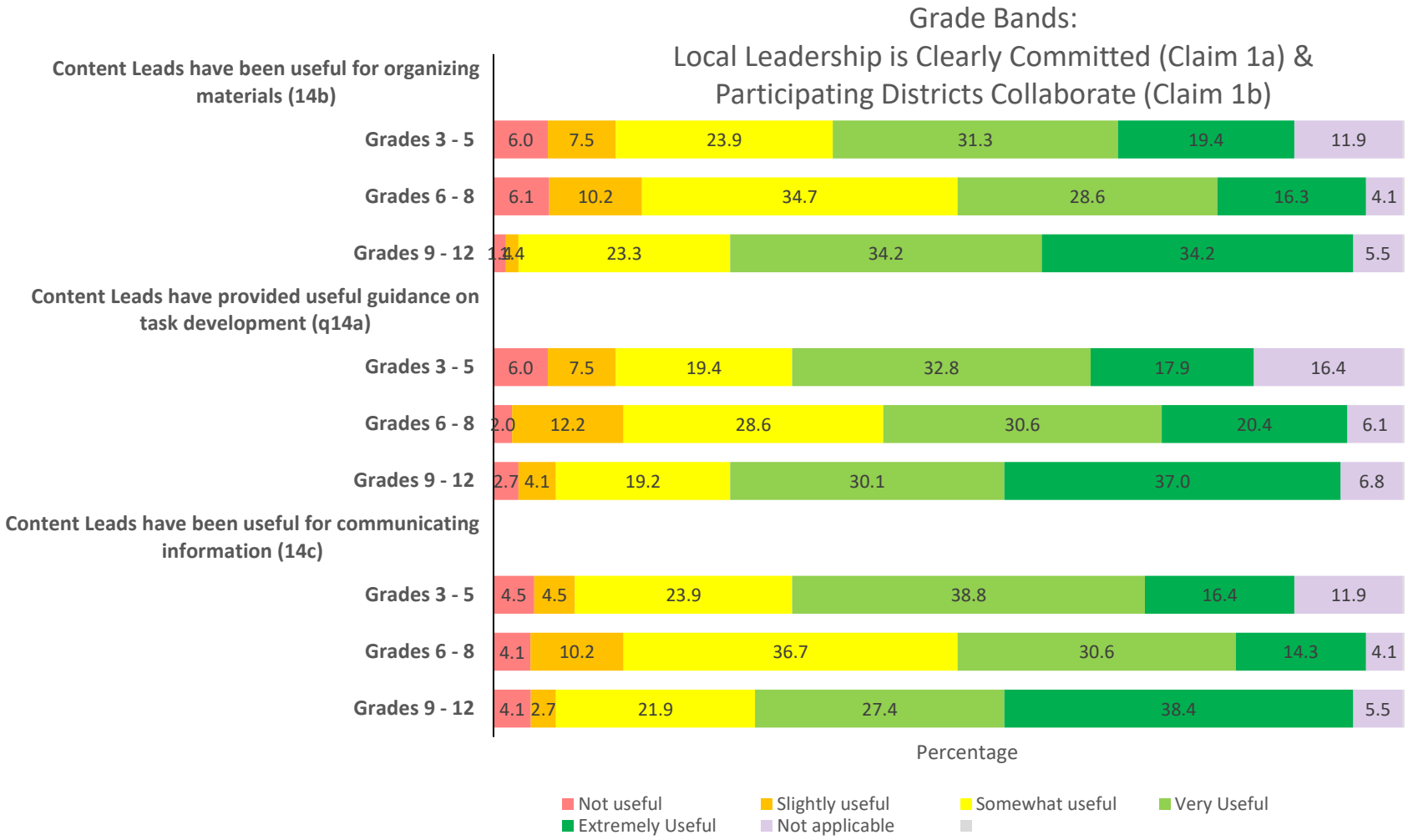


Figure 26. Grade bands comparison: Local leadership is clearly committed & participating districts collaborate (q14).

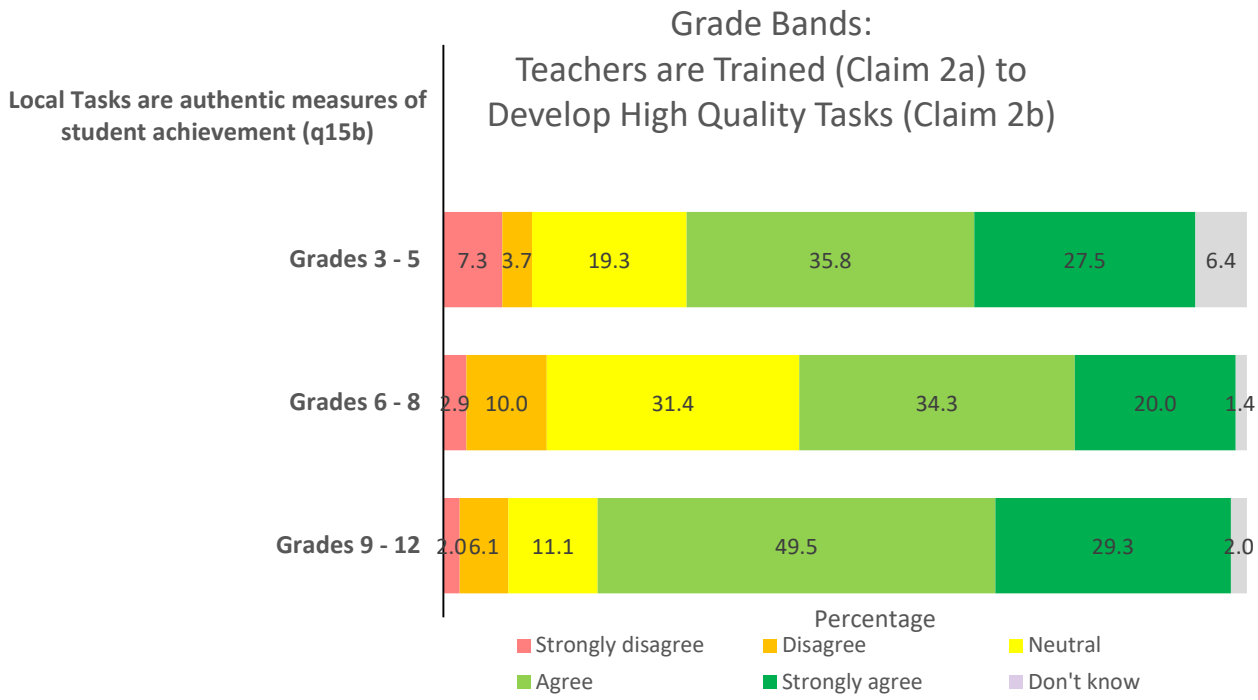


Figure 27. Grade bands comparison: Teachers are trained to develop high quality tasks (q15b).

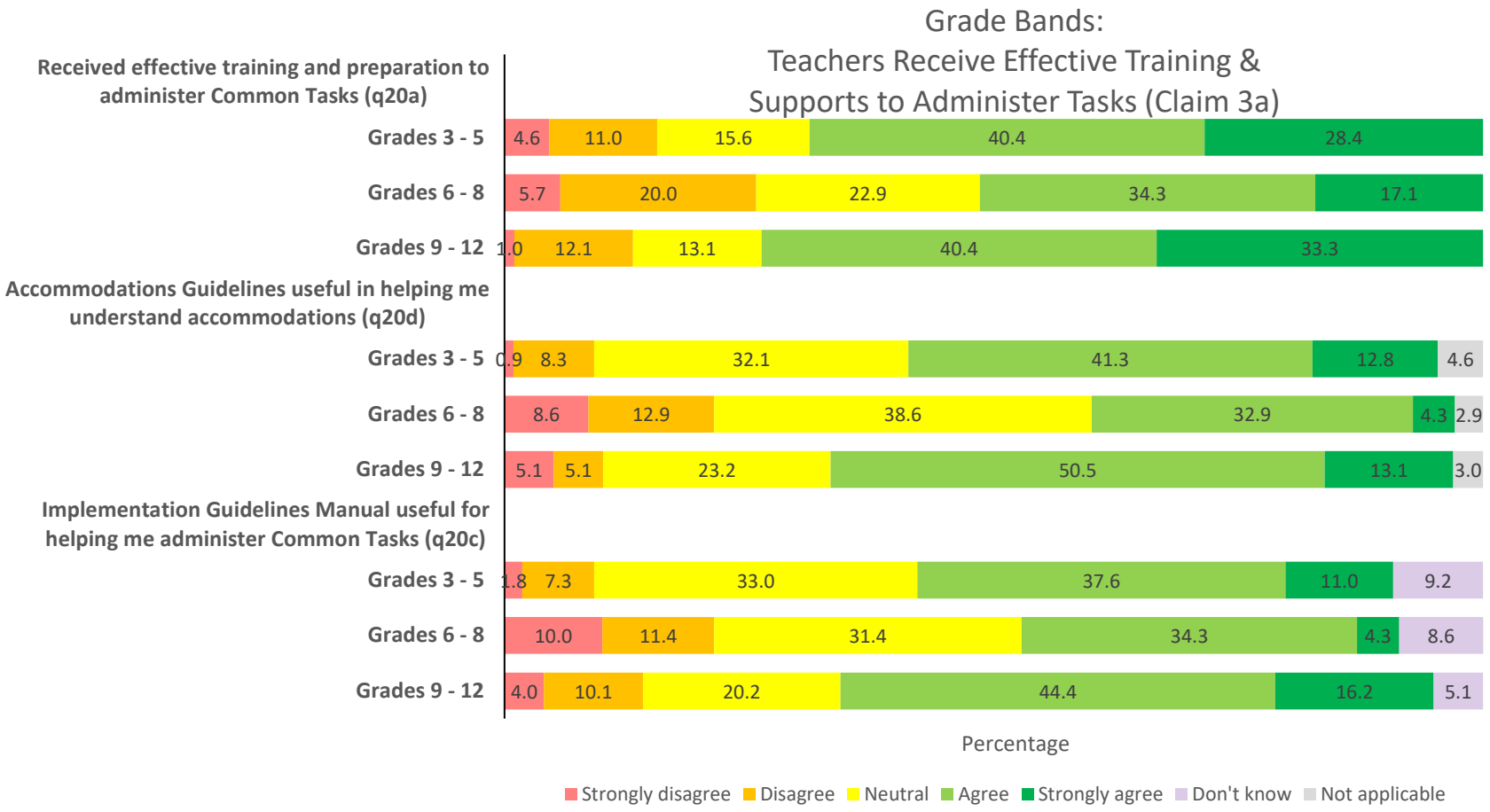


Figure 28. Grade bands comparison: Teachers receive effective training and supports to administer tasks (q20).

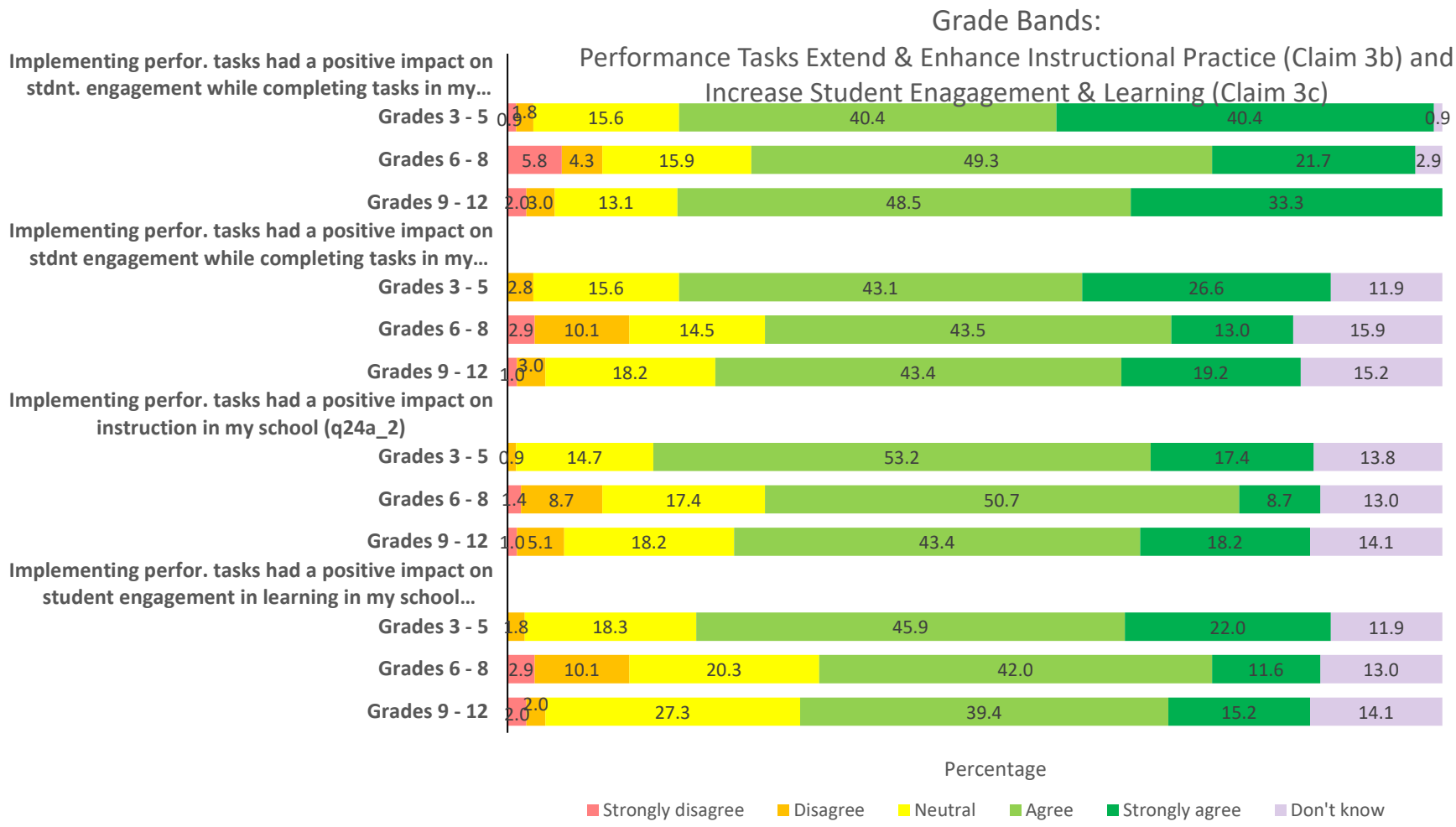


Figure 29. Grade bands comparison: Performance tasks extend & enhance instructional practice and increase student engagement & learning (q24).

Content Areas

Finally, differences in responses among participating PACE teachers teaching ELA, mathematics, and science were investigated. However, the survey was not designed to specifically address impressions for content area-specific performance tasks. For example, teachers were asked about the impact of implementing performance tasks on instructional practice, not about the impact of implementing *ELA* performance tasks on instructional practice. Consequently, teachers' responses to the background question, "indicate the content area(s) for which you are responsible for administering NH PACE Common Tasks" were used to disaggregate results. We analyzed these responses in two ways.

First, no substantive differences emerged when results were disaggregated by the content area the teacher selected on this background question. The problem with this approach, however, is that most elementary school teachers teach all three content areas. Consequently, there was considerable overlap in the data that were compared, possible suppressing real differences.

To help mitigate this, in a second analysis we included only those teachers that exclusively selected ELA, mathematics, or science. Again, no substantive differences emerged. However, this approach reduced the sample of respondents because nearly all elementary school teachers were excluded. Two teachers who responded to the open-ended comment at the end of the PACE Teachers Survey indicated that they had trouble answering some of the questions on the survey because their experiences were different for ELA and mathematics.

We have clear indications of content area differences through teacher focus groups and task development observations. These are discussed elsewhere in this report. However, we could not disentangle differences across content areas through this survey.

PACE Teachers Survey Results Compared to Other Teachers Survey Results

Eleven survey items were common to both the PACE Teachers Survey and the Other Teachers Survey. A comparison of results on those items indicates that Other Teachers responded similarly to the PACE Teachers. This suggests that PACE has had a positive impact school-wide.

Of the common items, there was only one item for which the non-PACE teachers were considerably more likely to select the "Don't know" response. Nearly 18% of the non-PACE teachers selected "Don't know" when asked about the opinion held by other teachers at their school regarding PACE, whereas only 6% of PACE teachers selected "Don't know." However, for all of the other common items, the percentages of "Don't know" and "Not applicable" responses were similar for both PACE teachers and non-PACE teachers.

The item mean ratings (with "Don't know" and "Not applicable" responses omitted) were slightly *higher* for the non-PACE teachers than for the PACE teachers. The magnitudes of the effect sizes were all relatively small (nothing greater than d = .35), but the trend was consistent across all of the items (see Table 11). This provides further support for the notion that PACE appears to be having positive impact school-wide.

Table 11. Mean Comparisons of Items Common to PACE Teachers Survey and Other Teachers Survey

Common Items[a]	Claim Addressed by Item	PACE Teachers		Other Teachers		Cohen's *d*
		Mean	S.D.	Mean	S.D.	
My school's administration is supportive of the PACE initiative.[b]	1a	4.17	0.87	4.42	0.80	-0.31
Opinion of teachers at my school on PACE.[c]	1a	3.29	0.89	3.37	0.86	-0.09
Extent to which competency-based education is integrated into my instruction.[d]	3b	3.81	0.83	3.89	0.88	-0.09
Implementing performance tasks has had a positive impact on instructional practice *in my classroom*. [b]	3b	4.09	0.82	4.21	0.75	-0.15
Implementing performance tasks has had a positive impact on instructional practice *in my school.* [b]	3c	4.06	0.92	4.17	0.83	-0.13
Implementing performance tasks has had a positive impact on student engagement while completing tasks, *in my classroom.* [b]	3c	3.96	0.87	4.13	0.81	-0.20
Implementing performance tasks has had a positive impact on student engagement while completing tasks, *in my school.* [b]	3b	3.88	0.79	4.09	0.79	-0.26
Implementing performance tasks has had a positive impact on student engagement in learning overall, *in my classroom.* [b]	3c	3.93	0.87	4.07	0.78	-0.17
Implementing performance tasks has had a positive impact on student engagement in learning overall, *in my school.* [b]	3c	3.82	0.86	4.01	0.81	-0.23
I use performance tasks for instructional purposes.[e]	3a	2.50	0.89	2.86	1.23	-0.34
I use performance tasks for assessment purposes.[e]	3a	2.19	0.67	2.50	1.06	-0.35

[a]Some of the item stems are abbreviated in Table 3. See full item stems in Appendix A.
[b]Items rated on 5-pt agreement scale where 1 = Strongly disagree and 5 = Strongly agree.
[c] Item rated on a 5-pt opinion scale where 1 = All have a favorable opinion; 2 = Most have a favorable opinion; 3 = Some have a favorable opinion; 4 = Few have a favorable opinion; 5 = None have a favorable opinion (this item is reverse scored).
[d] Item rated on 5-pt extent scale where 1 = To no extent and 5 = To a very great extent.
[e] Items rated on a 5-pt frequency scale where 1 = Never; 2 = Couple times a month or less; 3 = Approximately once each week; 4 = 2 to 3 times each week; 5 = Nearly every day.

Summary of Findings by Evaluation Goal

The following summary of findings is based on data collection activities conducted over the course of the April 2016 –February 2017 evaluation period, including activities detailed in this report as well as HumRRO's June, September, and Formative Evaluation: Final Reports. Findings are summarized below by the study's three evaluation goals. We have not repeated all of the findings detailed in our previous reports, but instead provide a brief summary along with any additional examples of findings from recent data collection efforts.

Evaluation Goal 1: Refine and Validate the PACE Accountability Program's Theory of Action

Upon contract award, HumRRO reviewed PACE materials and ensured that the graphical TOA supplied in our proposal to guide the evaluation was accurate. During the project kick-off meeting, we reviewed and discussed the draft TOA with the PACE Leadership team. The leadership team indicated that the draft TOA provided a useful and accurate framework for evaluating the PACE accountability program. No edits or revisions were requested at that time.

We reviewed the draft TOA with each of the eight PACE District Leads during our spring 2016 telephone interviews and with the new, ninth PACE District Lead in fall 2016. Although all individuals indicated that the PACE Program TOA is consistent with their goals for PACE in their district, the spring 2016 discussions resulted in two revisions to the TOA:

1. Two PACE District Leads noted that the process is continuous. For example, as new teachers enter the system they must be acclimated to and trained on PACE. Consequently, to better reflect the continuous nature of PACE, a loop has been added to the TOA in Figure 1.
2. Several of the PACE District Leads commented that the size of the district/school has an impact on the goals and claims in the TOA. For example, the PACE District Leads in smaller districts explained that due to resource constraints (e.g., fewer teachers and supporting staff) implementing PACE as intended is a challenge. On the other hand, a PACE District Lead in a large school district explained that it is difficult to effectively inform and train all their teachers on PACE. Therefore, to reflect the impact of district/school size on PACE, a "Contextual Factors" moderating variable has also been added to the PACE Program TOA in Figure 1.

Evaluation Goal 2: Provide Formative Feedback on Key Success Criteria

The findings for evaluation goal 2 are organized by the interim goals and claims in the TOA. During the course of this formative evaluation, we (a) attended numerous PACE events and meetings (both in-person and via teleconference); (b) conducted phone interviews with all nine PACE district leads; (c) visited each of the Tier 1 districts and conducted focus groups and interviews with students, teachers, administrators, and parents; (d) administered two on-line surveys—one for teachers administering NH PACE common tasks and one for other teachers in PACE Tier 1 schools; and (e) reviewed student data from the NH PACE common tasks. The evidence we gathered from these data collection efforts were mapped onto the claims from the TOA in order to help build the validity argument for the PACE accountability program. Where we find evidence of threats to the validity of the claims, we provide recommendations to help mitigate those threats (see the "Recommendations" section of this report).

TOA Interim Goal 1: Stakeholders Are Committed to PACE

Claim 1a: Local Leadership Is Clearly Committed

Over the course of this formative evaluation, evidence consistently emerged to support the claim that local leaders are clearly committed to PACE. We inquired about commitment during teacher focus groups, administrator focus groups, interviews with PACE district leaders, and surveys of PACE teachers and non-PACE teachers (i.e., those not responsible for administering the NH PACE common tasks). Across all these data collection efforts, most teachers indicated that their school administrators are supportive of PACE, including providing guidance and resources for implementing the NH PACE tasks, answering questions, and being knowledgeable. School administrators, in turn, reported that district leadership is supportive. Most teachers also reported that the teachers at their school effectively collaborate with one another on implementing PACE.

Many teachers and administrators noted that communications about PACE have improved since the first year of the PACE pilot. Several teachers mentioned that they are now able to contact a PACE task development expert directly and that they appreciate having access to this resource. Another commonly mentioned improvement was that information is more organized and is shared earlier, allowing school staff time to plan ahead for meeting PACE requirements. For example, teachers in the Concord school district explained that in the 2015–16 school year, many PACE tasks were administered at the end of the school year which caused a rushed and stressed timeline. This year, however, more advanced planning is occurring and PACE tasks are being administered as they fit the curriculum during the school year.

One ongoing source of remaining tension is the amount of time PACE requires, including time away from the classroom for task development and calibration sessions. PACE Leadership is aware of this concern and has been making efforts to address it. For example, a task development session was scheduled on a Saturday in September 2016 so educators could participate without sacrificing classroom time. In addition, appointing Content Leads and Teacher Representatives has served to limit the amount of time other teachers must spend outside the classroom on cross-district PACE activities. Evidence from recent district visits and from the PACE Teachers Survey indicate that these efforts were appreciated and useful. Nonetheless, concerns still exist regarding the time required by PACE. One of the questions on the PACE Teachers Survey asked teachers whether the time and effort required by the PACE initiative were worth the benefits that they have experienced and/or seen. More positive responses than negative responses were received; however, nearly a fourth of respondents disagreed with the statement. Disagreement tended to be stronger for (a) teachers from the White Mountains School District, (b) teachers who switched schools or districts since 2014–15, (c) teachers who did not participate in cross-district collaborations, and (d) middle school teachers. Consequently, overall, the findings across the data collection efforts indicate that the majority of stakeholders were clearly committed to PACE, however, there were some, albeit a minority, who reported that PACE was not worth the time and effort it required.

Claim 1b: Participating Districts Collaborate with One Another

A theme that emerged across the district interviews and focus groups with teachers and administrators was that teachers value and enjoy the opportunity for cross-district collaboration. They often referred to it as beneficial for their professional growth, and for aligning instruction and PACE tasks with other districts. Furthermore, findings from the PACE Teachers Survey indicate that the majority of teachers reported that collaborations with teachers from other

districts were useful for task development and calibrating scoring of student work. Teachers who responded to the PACE Teachers Survey reported that, aside from face-to-face communication, LibGuides were the most effective resource/method for facilitating cross-district collaborations. The majority of teachers who participated in cross-district collaborations also rated the guidance and support provided by the Content Leads as very useful or extremely useful. Collectively, these findings support the validity of the claim that participating Tier 1 districts effectively collaborate with one another.

There was, however, some evidence of potential threats to the validity of claim 1b. We conducted brief focus groups in the May 2016 Task Development Session during which teachers identified the infrequency of opportunities to meet, the fact that they use different curriculum programs, and the need to bring new member districts up to speed, as barriers to effective cross-district collaboration. Subsequently, PACE Leadership identified Content Leads to help with these collaborations. During the fall 2016 visits to Tier 1 districts, teachers from two districts and one school administrator mentioned differences in individual expertise and differences in local curricula as challenges for cross-district collaborations on task development. Most recently, during the December 2016 task development session, teachers were observed working together effectively, including identifying behavioral norms to guide their collaborative efforts, with emphasis placed on clear, open, and respectful communication.

PACE Leadership should monitor whether the addition of Content Leads and Teacher Representatives continues to improve some of the earlier noted barriers to cross-district collaborations. In addition, the PACE Leadership team may want to give special attention to (a) teachers who switched schools and/or districts since 2014–15, (b) teachers from the White Mountains School District, and (b) middle school teachers, as these particular groups of teachers tended to express somewhat less favorable opinions of cross-district collaborations on the PACE Teachers Survey. Also, during a December 2016 visit, teachers at Seacoast Charter School expressed concerns that because they are a small school with limited staffing resources, they were not always able to be involved in all stages of task development (e.g., initial brainstorming of tasks). They reported feeling that their involvement was limited.

TOA Interim Goal 2: Assessments Are Based on Sound Test Design Principles

Claim 2a: Teachers Developing Performance Assessments Are Trained On/ Knowledgeable of the Joint Standards[16] for Test Development

Although references to the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 2014) are not rampant in PACE, this formative evaluation focused on identifying evidence of effective assessment practices, including concepts of reliability and validity. Throughout the various data collection efforts there was consistent evidence of the PACE teachers' assessment literacy. For example, during the first on-site data collection—observations of ELA and science task development sessions in May 2016—the teachers exhibited a great deal of assessment literacy. While they did not talk about their tasks in formal psychometric terms, they were concerned about ensuring sufficiently standardized administration, guarding against score contamination, maintaining consistency between the tasks and their intended measurement constructs, how consistently and accurately the tasks could be scored with the rubrics, and guarding against construct irrelevant variance. They

[16] *We understand that the PACE stakeholders are not test design experts and, therefore, that the AERA, APA, & NCME Standards are not firsthand knowledge for this audience. Consequently, our discussion with these stakeholders referred more generally to "high-quality assessment."*

discussed fairness, opportunity to learn, and accommodations and supports for students who need them. The teachers struggled with these issues in much the same manner as testing professionals. Furthermore, in later task development sessions, the teachers and Content Leads used terminology that reflected a deep understanding of developing high quality tasks. For example, they asked questions such as, "What are the big competencies?" "What skills do students need to have for that competency?" and "What evidence shows mastery of that competency?"

More recently, during a December 2016 visit to the Seacoast Charter School, teachers reported improvements in their understanding of task development, including writing clear task instructions and developing scoring rubrics with distinguishable and achievable performance levels. Also entailed in the theory of action for PACE is that teachers apply what they learn from developing high quality *common* tasks to the development of high quality *local* tasks. Findings from the PACE Teachers Survey reveal that the majority of teachers reported that they were able to apply what they learned from the NH PACE common performance tasks to developing higher quality local performance tasks (although it should be noted that teachers from the White Mountains School District and teachers who did not participate in cross-district collaborations rated this less favorably).

These are a few examples illustrating consistent evidence of teachers' knowledge of effective assessment design principles. Overall, the evidence supports the validity of the claim that teachers who develop performance assessments are trained on and knowledgeable about principles of quality test development.

Claim 2b: Performance Assessments Adhere to the Joint Standards, Including Ensuring Equity

The PACE Teachers Survey asked teachers about the authenticity of PACE tasks as measures of student achievement. The majority of teachers who responded to the survey agreed that "NH PACE common performance tasks are authentic measures of my students' achievement" and that "Locally developed performance tasks are authentic measures of my students' achievement." This evidence supports the claim that performance assessments adhere to the *Joint Standards*. It is important to recall that teachers who switched schools or districts since 2014–15 and teachers who did not participate in cross-district collaborations tended to rate these items less favorably than other teachers. Additionally, middle school teachers also provided lower ratings for the authenticity of locally developed tasks.

The PACE Teachers Survey also asked teachers to rate their level of agreement with the statement, "NH PACE Common Performance Tasks are more accessible to a greater range of student learning needs (e.g., students with disabilities, English language learners) than traditional standardized tests." This survey item addresses the "ensuring equity" aspect of Claim 2b. The majority of teachers agreed with this statement, although nearly a fourth disagreed. Teachers' open-ended comments on the survey help to provide some additional, contextual information. The most common theme mentioned in teachers' open-ended comments were concerns related to the validity of PACE. It's important to note that even though this was the most common theme, it was only mentioned by about a third of the teachers who provided an open-ended comment, which represents just over a tenth of all the teachers who responded to the survey. Within that broad theme, several teachers mentioned specific concerns about the accessibility of the common tasks. Contextual detail provided in teachers' open-ended comments indicated that some teachers believed that the reading and writing demands for the common tasks were too high and, therefore, a barrier to accessibility for many students. This

concern was also raised by some teachers participating in a spring 2016 focus group at Sanborn Regional and at a winter/fall 2016 focus group at the Pittsfield school district.

Several additional sources of information also help to inform Claim 2b. First, during our initial site visit to the ELA and science task development sessions in May 2016, the majority of teachers who participated in focus groups at those sessions indicated that the PACE tasks are more accessible to students with a greater range of learning needs. They noted that the PACE tasks are more authentic and they allow students to respond in the same way that they do in the classroom. Facilitators from both the ELA and science sessions also commented that PACE tasks offer more pathways for students with disabilities (SWD) to access the tasks, whether through testing students when they are ready (ELA facilitators) or by engaging them through group work on performance tasks (science facilitators). There was, however, a subset of teachers from both the ELA session and the science session who expressed some concern about this topic—both sides' concerns were related to the notion of ambiguity. A few teachers in the ELA session felt that there was a lack of clarity about what is allowed and what is not allowed (e.g., can students discuss reading passages before responding to a writing assessment?), and some science teachers worried that scaffolding for SWD is not specifically addressed in their tasks. Similar sentiments were echoed in later data collection activities. That is, many noted that performance tasks allow multiple access points for SWD, but that there is room for improved clarity around accommodations and scaffolding. This concern has consistently been voiced by a minority of teachers, but persists.

Collectively, the majority of the evidence supports the claim that performance assessments adhere to the *Joint Standards*, including ensuring equity. However, there is some concern regarding a lack of clarity surrounding accommodations and scaffolding, and that the reading and writing demands of the common tasks may be too high for some students to access the tasks.

TOA Interim Goal 3: Performance Assessments Are Successfully Implemented

Claim 3a: Teachers Receive Effective Training and Supports to Administer the Performance Assessments with Fidelity

The first set of visits to school districts occurred during spring 2016 and included interviews and focus groups with teachers and administrators from Sanborn Regional, Souhegan, and Rochester school districts. Overall, findings from those data collection activities indicated that teachers felt they received adequate training and supports to administer the PACE tasks as intended, although teachers from the Rochester School District expressed interest in more formal PACE training. The second set of site visits to Monroe, Epping, Pittsfield, Concord, and the Seacoast Charter School occurred during fall/winter 2016. The site visit to Monroe revealed that the Monroe teachers also felt well-prepared to administer the common performance tasks, citing the quality of training and the specificity of directions. Epping teachers praised the common tasks (especially science) but some noted that tasks are interpreted in different ways by teachers who are not part of the development process. Concord teachers indicated that they received sufficient training and support, and that help was available from PACE Content Leads, which they frequently reached out to when questions arose. Seacoast teachers reported that they felt they received sufficient training regarding task development, scoring, and calibration, but they noted that there was less training on administration. Some teachers at Pittsfield expressed challenges with training such that they were left to "figure it out," but they indicated that they found the resources on LibGuide to be valuable. They recommended that training

videos to show correct task administration would be helpful to ensure that all teachers administer common tasks consistently.

The PACE Teachers Survey also provides useful information to inform Claim 3a. More than three-fourths of the teachers reported that their school's administration provides them with the resources and supports they need to effectively implement the NH PACE common performance tasks. More than two-thirds reported that they received adequate training and preparation to effectively administer the common tasks. More than half reported that the NH PACE Implementations Guideline Manual and the NH PACE Accommodations Guidelines were useful resources in helping them understand how to effectively administer common tasks. Also, nearly two-thirds of the teachers reported that, based on the scaffolding rules, they understood the amount of scaffolding they can employ with common tasks, although 21% disagreed, which relative to other items on the survey, was a comparatively high level of disagreement. (It should be noted that compared to other teachers, teachers from the White Mountains School District, teachers who switched schools or districts since 2014–15, teachers who did not participate in cross-district collaborations, and middle school teachers tended to rate the effectiveness of the training and supports they received for administering common tasks less favorably).

The majority of the evidence supports the claim that teachers receive effective training and supports to administer performance tasks with fidelity, although there is evidence to suggest that some teachers feel that training and supports for task administration could be improved. Additional clarity and guidance on the amount of scaffolding that can be employed with common tasks may be one specific area to target.

Claim 3b: Implementing the Performance Assessments as Intended Enhances and Extends Desired Instructional Practices

Most of the evidence collected across the data collection activities supports the claim that implementing performance assessments as intended enhances and extends desired instructional practices. Teachers across participating Tier 1 districts expressed that implementing performance tasks has had a positive impact on their instruction. They commonly mentioned that PACE has had a positive impact on increasing the DOK at which they teach and gives them real-time feedback that they can use to make "on-the-spot" adjustments to their instruction to better meet the needs of their students. Teachers at both Monroe and Seacoast Charter School also expressed that implementing PACE tasks has contributed to more coherence and focus on the content of their instruction. Teachers from Concord noted that because PACE is embedded into their curriculum, it doesn't take away from their planning time and instruction is not put on hold during assessment. Some Epping teachers noted that their scope and sequence must be modified to accommodate PACE tasks, and students who are not on level academically must work out of level for a month or two. Teachers from the Pittsfield School District noted the need to spend instructional time familiarizing students with the format of the tasks and the need to re-organize their schedule to ensure relevant lessons and competencies were presented in time for the PACE task. Some teachers echoed similar concerns in their open-ended comments on the PACE Teachers Survey.

The teacher surveys provided a broader view of teachers' perceptions than did brief focus groups with a subset of teachers. Overall, the findings from the PACE Teachers Survey and the Other Teachers Survey indicate that both "PACE Teachers" and non-PACE Teachers agreed that implementing performance tasks has had a positive impact on their instructional practice in their classroom, such that instruction occurs at a higher DOK. This was among the highest rated items on the PACE Teachers Survey and the Other Teachers Survey. Teachers were also

asked about PACE's impact on instruction at the school level. Results were also very favorable for PACE Teachers and non-PACE teachers. Teachers from White Mountains and teachers who switched schools/districts did *not* rate these items substantively lower than other teachers. Teachers who did not participate in cross-district collaborations and middle school teachers did tend to rate these items less favorably.

An assumption of Claim 3b is that competency-based education is integrated into teachers' instruction. Most teachers who responded to the PACE Teachers Survey and the Other Teachers Survey indicated that competency-based education is integrated into their instruction to a great or very great extent, thereby lending further support to Claim 3b. Although teachers who did not participate in cross-district collaborations tended to rate this item less favorably than other teachers, teachers from the White Mountains school district were *more likely* than teachers in other Tier 1 districts to indicate that competency-based education is integrated into their instruction to a great or very great extent.)

Another assumption underlying Claim 3b is that teachers use performance tasks for instructional purposes (i.e., not just exclusively for assessment purposes). Findings from the focus groups with teachers help to substantiate this assumption. The findings from the surveys lend further credence to this assumption. Only 8% of teachers who responded to the PACE Teachers survey indicated that they "never" use performance tasks for instructional purposes; the majority (52%) indicated that they use performance tasks for instructional purposes once or twice a month, and the remaining 40% reported that they use performance tasks for instructional purposes more than a couple of times a month. Even the non-PACE teachers (i.e., those not responsible for administering NH PACE common performance tasks) overwhelmingly reported that they use performance tasks for instructional purposes. Only 9% of the non-PACE teachers reported that they "never" use performance tasks for instructional purposes. The majority (42%) of the non-PACE teachers indicated they use performance tasks for instructional purposes once or twice a month, and the remaining 49% reported that they use performance tasks for instructional purposes more than a couple of times a month. This suggests that PACE is consistent with school-wide instructional practices (i.e., not just among those teachers administering the NH PACE common performance tasks).

Claim 3c: Student Engagement and Student Learning Increases/Deepens When Performance Assessments Are Implemented as Intended

The findings from the data collection activities provide support for this claim. Corroborating evidence to support Claim 3c came from focus groups and interviews with key stakeholders, classroom observations, and teacher surveys.

The site visits to Tier 1 schools included focus groups with teachers, students, administrators and parents. During the teacher focus groups, across all districts, teachers and school administrators commonly reported that PACE has had a positive impact on student engagement. For example, several teachers noted that students are more involved with/invested in the work they are doing for the PACE tasks, and that both teachers and students see the PACE tasks as learning, not just an assessment. The science tasks, in particular, were discussed as lending themselves to collaboration, which facilitates student engagement. Some school leaders and teachers also noted that often students do not even realize that they are taking a "PACE task." Students just think it is part of what they do on a normal basis. This was noted by some as helping to alleviate the test-related anxiety that is commonly associated with more traditional, standardized tests.

When students were asked what they liked most about PACE many noted that they like how PACE requires more application of real world skills and how they can demonstrate their knowledge in multiple ways. Many students also reported that they found the PACE tasks more challenging and also more interesting than the "bubble tests" and the "computer tests." Some students did note that there was not enough time to complete the PACE tasks, especially those at the middle school and high school grades; they explained that breaking up the assessment over the course of a week within a normal class period interrupted their thinking. This concern was also noted by some teachers during the site visits. Students expressed some mixed feelings about the collaborative component to PACE tasks. Some students thought it was beneficial to be able to work with others and use their strengths where they might be weaker, but others thought that if the group dynamic was not supportive, the collaboration could be more harmful than beneficial.

During parent focus groups, parents noted that PACE tasks encourage a deeper level of understanding than a traditional multiple choice test, and said that the preparatory work, and the task itself, causes students to retain their learning longer. Multiple parents noted that the PACE tasks train students to self-critique and have ownership of their learning. Some parents disagreed as to whether students were more engaged in PACE tasks than other classroom assignments or tests. Some felt that the application piece makes the learning deeper, while others suggested that the tasks may be more entertaining and interesting but not necessarily deeper.

There were few opportunities to observe administration of operational common tasks, although there were several opportunities to observe administration of other, locally developed tasks. Classroom observations at the Concord school district showed students involved and engaged in learning. Teachers used hands-on and multiple methods when interacting with students (e.g., smart boards, iPads, group work). In Sanborn, students were observed highly engaged in their tasks, and when asked, could clearly communicate what they were doing and why. Souhegan students participating in a PACE operational task were very engaged and focused, and appeared to follow the task instructions and work diligently to demonstrate their learning. Students appeared comfortable implementing unique approaches to their experimental design and confidently applied their prior knowledge to completing the task.

Finally, the results from the teacher surveys indicate that both PACE Teachers and non-PACE Teachers (i.e., those not administering NH PACE common tasks) overwhelmingly reported that implementing performance tasks has had a positive impact on student engagement while completing tasks *and* on student engagement in learning overall, particularly at the classroom level. Schoolwide impact was also rated favorably by PACE teachers and non-PACE teachers alike. Disaggregated survey analyses indicate that teachers who have not participated in cross-district collaboration and middle school teachers tended to have less favorable perspectives on the impact of implementing performance tasks on student engagement and learning, although, overall, their ratings were still quite positive.

Collectively, these findings indicate support for the claim that student engagement and student learning increases/deepens when performance assessments are implemented as intended.

TOA Interim Goal 4: Scores Are Accurate and Reliable

We found considerable evidence for the claim that students' scores and annual determinations are accurate and reliable. Scorers were effectively trained and PACE tasks were double scored. The common task was used to equate among the districts and to evaluate scorer accuracy.

Claim 4a: Scorers Are Effectively Trained

The data collection efforts afforded opportunities for observing scoring calibration sessions, the PACE Tier 1 Summer Institute Body-of-Work (BOW) exercise, and surveying teachers about their perceptions of the effectiveness of scoring resources. Overall, the findings from these data collections support the claim that scorers are effectively trained.

Teachers participating in the calibration sessions reported that the calibration process for the common tasks has helped them with scoring at the local level. During the observed calibration activities, teachers were prepared, professional, and appeared to be engaged, focused, and working diligently. They each had identified potential anchor papers to discuss and were knowledgeable about the content and the rubrics. The PACE Tier 1 Summer Institute BOW exercise was a strong professional development opportunity for the select group of educators who attended; the training and hands-on work deepened their understanding of the PACE assessment system. Some participants did express a desire to have examples of student work that represented high scores. As a result of this request, high score examples are now provided in the data collection resources in the administration LibGuide.

The majority of teachers who responded to the PACE Teachers Survey agreed that "The scoring rubrics for the NH PACE common performance tasks are sufficiently clear and detailed to ensure that separate scorers scoring the same student work arrive at the same score" and "The scoring resources available on the LibGuide effectively explain how to score student work on the NH PACE Common Performance Tasks." However, approximately a quarter of the respondents disagreed that the scoring rubrics are sufficiently clear and detailed to ensure that separate scorers scoring the same student work arrive at the same score. Although this was a minority of respondents, this was a comparatively high percentage of disagreement compared to the other items on the survey. Some teachers also mentioned concerns about the clarity of the scoring rubrics in their open-ended comments on the survey. (Teachers from the White Mountains school district were more likely than other teachers to respond "Don't know" when asked about the clarity of scoring rubrics. Teachers who have switched schools or districts since 2014-15 were more likely than other teachers to provide less favorable ratings on effectiveness of the scoring resources available on the LibGuide.) Collectively, the findings across the data collections support the claim that scorers are effectively trained, although there is some evidence to suggest that some scoring rubrics may benefit from additional detail and clarity.

Claim 4b: Scorers Attain Successful Rates of Inter-rater Agreement and Reliability

The PACE common task is a key component for ensuring the success and viability of the overall assessment system. While a specific level of inter-rater agreement is not a requirement for local scoring of student responses, during the PACE Tier 1 Summer Institute, participants conducted consensus scoring of a sample of student work. Each pair of scorers represented two districts and the student responses they reviewed came from a third district. They did not have access to the original teacher scores and they generally had little difficulty reaching consensus.

Subsequent to the consensus scoring meeting, the scores from the central scoring group are compared with the scores from the district. If there is poor agreement between the district results compared to the consensus scored results, the scores on the common tasks are adjusted to account for the discrepancy. If the differences between adjusted common task performance is substantially different from local task performance, it may also signal a district level scoring bias. If such a difference is discovered, scorers can be retrained on a district by district basis.

Within-district inter-rater reliability is computed by the Center for Assessment. They determine whether a teacher scores more leniently or strictly by comparing the teachers' scores on the common task to the consensus scores on that task. The index they use for this purpose is a "deviance" index, which describes how far from the consensus scored papers an individual teacher scores (averaged across students). Several flags for potentially inconsistent scoring have been established, but scoring for 2015–16 was quite consistent. While there were minor differences between subjects and by district, scoring for PACE common tasks by teachers was largely verified as accurate and consistent during consensus scoring.

The Center for Assessment also computes within-district rater agreement statistics (e.g. % exact agreement, % adjacent agreement) and Cohen's Kappa statistics for a sample of the double-scored common tasks (Evans & Lyons, 2016). Pairs of raters had exact agreement rates of between approximately 60 and 85%. There were substantial differences by grade, subject, dimension, and by district, but nearly all districts achieved greater than 60% exact agreement rates across all grade subjects. Kappa statistics indicate moderate to substantial agreement of ratings across all grades and subjects as well.

Samples of local tasks are also double scored. Teachers examine the results, but formal reliability statistics are not monitored during active scoring. Students' scores on the local tasks represent their work over the course of the year. They might be compared with more typical end-of-unit test scores. Unlike typical end-of-unit tests, students receive rubrics along with their PACE task instructions. This allows them to self-monitor as they work. If, at the end of the task the teacher score is different from the students' expectations, they can discuss the differences with the teacher. This provides the teacher with a quality check on the rubric and gives the student an opportunity to understand how to interpret and use the rubric to achieve the score they desire. Also, parents noted that the rubrics provide information to facilitate a discussion with their children about their performance on the task. The rubrics provide clear expectations for the students who use them, which improves the validity of the scores. The feedback teachers receive from the double scoring and from their interactions with students helps improve their locally developed tasks and rubrics to achieve better reliability.

More than two-thirds of the teachers responding to the PACE Teacher Survey reported that within-district double scoring is effective for ensuring the reliability of scores on the NH PACE common performance tasks. Teachers who did not participate in cross-district collaborations were less likely to agree with this item than teachers who participated in cross-district collaborations but, overall, their ratings were still quite positive. Most of the teachers from the White Mountains School District had not yet had an opportunity to score student responses and responded "Don't know" when asked about the effectiveness of within-district double scoring for ensuring reliability of scores.

Overarching TOA Claims: Negative Consequences and Construct Irrelevant Variance Are Minimized

PACE was implemented, in part, to reduce perceived negative consequences associated with large-scale end-of-year standardized testing. PACE was designed to stave off reductions in the depth of learning of students, to promote critical thinking, and to integrate curriculum, instruction, and assessment into a cohesive system of education. We have discussed some of the ways that PACE has succeeded in reducing the negative consequences that existed in New Hampshire schools, but it is also important to recognize potential negative consequences of PACE and to guard against them.

In previous reports, we indicated that the most prevalent challenge cited by teachers was finding time to create, administer, and score the performance tasks. Additionally, the time spent out of the classroom for task development was commonly identified as a challenge by teachers, administrators, and PACE District Leads. In recent months, teachers reported clearer communication regarding data requirements and timelines as a move in the right direction. Further, PACE Leadership has implemented a Content Lead role to limit the amount of time required by other teachers to create tasks, and a Teacher Lead role to provide a point person in districts without a Content Lead. Also, the addition of the Saturday task development meeting on September 24, 2016, provided an opportunity to further task development while limiting the amount of time teachers were out of the classroom.

Overall, the findings from the surveys indicate that both the PACE teachers and the non-PACE teachers have favorable opinions about the PACE pilot program. There were no items on either the PACE Teachers Survey or the Other Teachers Survey that received a mean rating below 3.00 (i.e., the mid-point on the Likert scale) or that received more unfavorable ratings (i.e., 1s or 2s on the Likert scale) than favorable ratings (i.e., 4s and 5s). Nonetheless, the survey results do provide insights into a negative consequence of PACE and also some potential sources of construct irrelevant variance.

The survey results provide additional evidence that many, albeit a minority of the population of PACE teachers, have concerns about the amount of time and effort required by PACE. Approximately a fourth of the PACE teachers disagreed or strongly disagreed that the time and effort required by the PACE initiative are worth the benefits they have seen and/or experienced. The addition of Content Leads appears to be a useful addition to help manage some of the PACE requirements, as most teachers provided favorable ratings on the usefulness of Content Leads. Additional efforts to further reduce the burden placed on teachers should continue to help reduce the number of teachers who do not feel that the benefits of PACE are worth the time and effort it requires.

One negative consequence was evident in middle school science. New Hampshire does not currently have a grade-by-grade curriculum for middle school science, but the common science tasks are grade specific. There is, therefore, some concern among educators that the tasks do not always match their curriculum. If, for example, one district teaches life sciences in grade 8, while another teaches physical sciences in grade 8, a common task in grade 8 related to life sciences could potentially disadvantage the latter district. The science tasks for middle school have been designed to address science and engineering principles and cross-cutting concepts, but they are not content free.

Teachers participating in an Algebra task development meeting demonstrated understanding of the implications of construct irrelevant variance. During a discussion of the technology requirements of a particular PACE task being developed, the group acknowledged the need to ensure that the task design did not yield scores that were more a reflection of technology skill than of mastery of Algebra competencies. The teacher survey results also suggest that construct irrelevant variance could be reduced by enhancing the clarity of the scoring rubrics as well as the resources available on LibGuide for scoring student work.

Finally, the survey results suggest that the accessibility of the common tasks for students with a greater range of learning needs (e.g., students with disabilities, English language learners) may benefit from further investigation, particularly regarding the reading and writing demands required by the common tasks.

Evaluation Goal 3: Capture the "Story" of PACE

PACE has lofty ambitions. Ideally, PACE will lead to an integrated competency based education system that is unbound by time in class, age, location where learning takes place, and other artificial methods of categorizing students. Instead, the system would focus on a core set of competencies and move students to the next phase of their education regardless of when, where, or how the student achieves those competencies. The system will incorporate a large number of ways for students to demonstrate the competencies and demonstration will take place in an on-demand way, where students can choose to take a performance event when they are ready, rather than according to a school calendar. Instruction would be more individualized and targeted toward the next competency the student needs to master. Such a system would reduce non-productive redundancy and allow students to learn at a much faster and more customized rate. Such a system would represent a dramatic shift from the traditional system of schooling.

PACE, as it is implemented currently, has taken steps toward this ideal. The PACE Tier I districts have begun identifying important competencies and they have designed performance tasks to measure those competencies. They have begun to build a bank of high-quality performance tasks that can be drawn on throughout a student's academic preparation. These tasks are made available to Tier 2 and 3 districts to be used as teaching tools, once they are available in the task bank after the administration year. They have moved toward a more integrated system of curriculum, instruction, and assessment, whereby assessment is being woven into all aspects of teaching and learning. Consideration of assessment when planning curricular sequence and planning lessons has increased among teachers since joining PACE. Students, even those who reported not liking PACE, described the tasks as complex and difficult while also being strong measures of their knowledge, skills, and abilities.

The scores generated from the PACE tasks are sufficiently reliable for their intended use and they are valid for uses beyond those that can be gained from more traditional end-of-year tests. Students reported understanding where they performed well and where they did not. PACE allows students the opportunity to redo their tasks (not for accountability purposes) once they have addressed the areas where they were not quite ready to demonstrate competency.

PACE has had a great deal of early success, but there is still a long road ahead if PACE is to realize all of its bold goals. First, PACE has to prove the program is sustainable. The program is relatively new and a few highly motivated districts have been instrumental in implementing the system. As new districts join PACE, there will be challenges. Getting new staff members oriented to such a complex and new way of educating students takes considerable time and effort. If the experienced teachers train the new ones, they will need time to do so. This time will be in addition to the time they spend implementing PACE in their own schools and classrooms. There may also be performance gaps between the experienced and newly joined districts. These issues, as well as potential changes in the political and economic climate in which PACE is being implemented will likely challenge PACE. The sustainability of PACE will rely on demonstrating that the benefits of PACE membership continue to outweigh the challenges. For this to happen PACE will require continuous feedback and improvement as the system expands.

The current PACE program has been very responsive to challenges and has improved based on feedback. For example, task development and piloting have been accelerated to make sure every task is sufficiently piloted and revised before it is used operationally. Communication regarding data collection, in-person meetings, and other important calendar-specific activities

have been improved and teachers have received this information earlier in the year. This helps teachers plan and makes the PACE program more readily implemented. PACE has begun to distribute minutes from Leads meetings as a means of ensuring common understanding of decisions and future plans. PACE has established Content Leads and Teacher Leads to limit the time teachers must spend outside their classrooms. All of these examples of program improvements resulted from PACE leadership responding to requests from teachers.

In addition to the improvements the PACE program has already made, more enhancements are in the offing. PACE Leadership plans to accelerate task development even more. The goal would be to allow pilot testing of the common tasks to begin in the fall semester if that is the most appropriate time in the curriculum to use them. This would allow a more genuine piloting of the tasks and provide data even earlier to facilitate review and revision of the tasks and rubrics. The PACE Leads are also discussing senior projects and senior exhibitions as a natural extension of this work. One of the monthly PACE Leads meetings was devoted to a presentation related to senior projects and exhibitions. The group decided that it was a sufficiently good idea to create a separate sub-group to explore ideas for implementing these new assessment components. This group is meeting on a monthly basis.

In addition to sustainability, the PACE program must also prove that it is scalable. New districts are joining PACE, but there does not seem to be an expectation that the program will eventually be state-wide. However, if PACE proves to be a substantially better program for educating students than the one that currently exists, it stands to reason that PACE should expand. PACE is currently adopted at the district level. This is, in part, because New Hampshire districts are extremely autonomous. It is, after all, the “Live Free or Die” state. Other states may not be structured similarly. Still, there is a great deal of preparation a district must do to become a Tier 1 PACE district. It would be difficult to suddenly implement PACE on a much broader scale because of the integrated nature of task development, teacher professional development, and collaboration. Getting a full state’s population of teachers to abruptly begin to effectively collaborate seems unlikely. In New Hampshire, PACE began with a few highly motivated districts and is expanding carefully. This model seems to be effective for a program like PACE, and if the program is transported outside New Hampshire, other states may want to adopt a similar implementation plan.

Recommendations

The recommendations in this section stem from the data collected during the course of the evaluation only. There is little literature that can be directly referenced and applied to a system like PACE. For that reason, there are no statements in the recommendations section that reference aspects of similar successful programs. We did not find systems that were both successful and sufficiently similar to PACE to make direct comparisons.

The recommendations also reflect that PACE is currently functioning largely as intended. The early success of PACE is well documented throughout this technical report. No broad or sweeping recommendations are indicated. The recommendations included here call for additional monitoring or minor improvements to current processes. As the system expands, more substantial changes may become necessary, but this evaluation does not indicate a need for major modifications at this time.

Throughout this evaluation we have provided formative feedback to PACE Leadership and noted ongoing improvements to the PACE program. In addition to improvements identified elsewhere, the documentation on the on-line LibGuide has expanded and been reorganized to be more accessible. Below we address progress regarding recommendations made in our most recent interim report (Becker et al., 2016c). We then provide a final list of recommendations as of the end of this evaluation.

Progress on Previous Recommendations

Most of the general recommendations proffered in our December 2016 report have been, or are in the process of being, addressed. These are examples of continuous process improvement as the PACE program matures. We list them here with updates from the current reporting period.

December 2016 Overarching Recommendation: PACE Leadership should continue the efforts that have helped PACE evolve into a strong and viable program for assessment and for improving instruction. Notably in the past two reporting periods, teachers and school administrators all pointed to communication as an area where PACE continues to improve. Teacher and school administrator feedback on the benefits of PACE was overwhelmingly positive. The teacher survey was administered to all teachers in Tier 1 schools and provided a more thorough window into the perceptions of PACE beyond the highly invested, key players, which revealed a consistent pattern of high ratings of PACE.

December 2016 General Recommendation 1: Monitor efforts to limit teacher time out of class to ensure it does not become a large problem. PACE Leadership took two steps to reduce the time required of teachers out of the classroom. A task development session in September was scheduled on a Saturday so it would not conflict with class time. Content Leads were identified, who conduct task development work, among other things, in between larger task development meetings. Surveys are conducted at the end of each of these events so educators have an opportunity to provide feedback, which can serve as one mechanism to monitor teacher perceptions regarding the amount of time they spend outside of their classrooms.

December 2016 General Recommendation 2: Ensure that pilot tasks and rubrics are sufficiently tested so any revisions prior to operational use are evidence-based. In the December 2016 Task Development days, PACE Leadership collected information on the pilot process and determined that 95 teachers will pilot the 17 2016–17 tasks by March 2017. In

March 2017 the tasks, rubrics, and supporting materials will be revised as warranted by the pilot experience, and the final pilot period will commence during the remainder of the school year. This process is substantially accelerated from the prior year, with an expanded number of teachers piloting the tasks.

December 2016 General Recommendation 3: Continue to look for ways to decrease the administrative burden placed on program participants. Two improvements in the 2016–17 school year were the full development of a school year calendar and identification of data expectations, such as collecting samples of student work throughout the school year. This will eliminate the retroactive work teachers conducted in the 2015–16 school year. Teachers responding to the PACE survey indicated that Content Leads provided useful guidance and support on PACE requirements, including organizing materials needed for task development (q14b). However, there still may be work to be done. Nearly a fourth of the PACE teachers reported that the time and effort required by PACE is not worth the benefits they have seen and/or experienced.

December 2016 General Recommendation 4: Monitor the effectiveness of recent efforts to improve communication. During the fall 2016 school visits, teachers and administrators noted that communication has improved. Further, most surveyed teachers agreed or strongly agreed that their school's administration communicates information to them about PACE requirements in an effective and timely manner.

December 2016 General Recommendation 5: Consider providing supplies or a supplies budget for PACE task materials. In the early days of the evaluation we heard some complaints from teachers that their out-of-pocket expenses to secure materials for PACE tasks were extensive. We did not hear similar concerns during our fall school visits. In fact, several teachers showed us their new tablets for students and expressed appreciation that their schools were providing for them well. In the teacher survey, the statement "My school's administration provides me with the resources and supports that I need to effectively implement the NH PACE Common Performance Tasks" had the highest mean rating on the survey. Over 80% of respondents agreed or strongly agreed with this statement.

December 2016 General Recommendation 6: Continue efforts to accelerate development of pilot tasks. As of this report, some pilot tasks for the 2016–17 school year are still under development. The addition of a Saturday task development meeting in September to "jump start" progress on pilot tasks was a step in the right direction. However, ideally and by design, these tasks would be available at the beginning of each school year to allow teachers to build them into the appropriate point in the curriculum of each class. Some content area/grade level teams are already beginning development of pilot tasks for the 2017—18 school year, which should position them to be ready by fall 2017. Availability of pilot tasks at the beginning of each school year will allow the scheduling of pilot activities to be more seamless.

Current Recommendations

We offer the following recommendations as this evaluation winds to a close. They are organized below within each Interim Goal.

Recommendations for Interim Goal 1: Stakeholders Are Committed to PACE

Recommendation 1: Monitor and Support District Engagement

PACE should regularly gauge local leadership support and target interventions when district leaders voice concerns or reduce their district's involvement with the program. PACE has done this for one district by helping support a PACE District Lead with experienced consultants. As the program expands, these checks and interventions should become more routinized to ensure that all districts maintain adequate support for the educators implementing the program.

Recommendation 2: Evaluate Effectiveness of Collaboration Methods

PACE has adjusted collaboration activities to address early concerns about the amount of time teachers must spend outside the classroom and communications issues. PACE should evaluate the effectiveness of the new collaboration methods. Although task development meetings with teachers from all Tier 1 districts were becoming unwieldy, one of the attributes teachers reported as positive was having direct input into the program. The more dispersed that input becomes, the less obvious individual teacher's input may be. If some teachers perceive the PACE program as coming from the outside rather than as a direct result of their own work, buy-in could suffer. Regular monitoring and adjustments can help safeguard against this potential issue.

Recommendations for Interim Goal 2: Assessments Are Based on Sound Test Design Principles

Recommendation 3: Consider Additional Training/Supports for Teachers Not Directly Involved In Task Development

As the percentage of PACE participants not directly involved in future common task development grows (either through including a smaller number of teachers in a meeting or by expanding into additional districts), the professional development and training stemming from those activities may need to be supplemented with additional training. Teachers routinely reported that the process of developing the common tasks greatly improved their own task development process and their approach to assessment. As the program expands, it will be important to maintain that benefit for all participants. One suggestion made by teachers in focus groups was to provide training videos showing appropriate and inappropriate task administrations as well as appropriate and inappropriate preparation activities.

Recommendation 4: Infuse Equity and Accommodations Training into PACE Activities

Despite quality documentation and training, teachers continued to report uncertainty regarding equity issues, especially for accommodating SWD. As the system expands and as attrition necessitates the inclusion of new teachers, it is important that these issues continue to be addressed to ensure both accessibility and validity. We suggest making equity and accommodations training part of the regular schedule of PACE activities.

Recommendation 5: Investigate the Impact of Reading/Writing Requirements on Accessibility

Investigate the impact of the reading and writing demands of the PACE tasks on accessibility and student performance. Several teachers indicated concerns that the reading and writing requirements for PACE were much higher than for traditional assessments. This can potentially

result in reduced test score validity, especially for SWD. This phenomenon occurs when the reading/writing load interferes with the measurement of the intended construct. If, for instance, we are interested in knowing whether student understand and can perform computations associated with a mathematics concept, including a long reading passage to set up the task might interfere with a student demonstrating her math abilities. We recommend examining score patterns among the PACE tasks, course grades, and performance on comparison measures (e.g., Smarter Balanced) for students with and without disabilities as one way to investigate whether the reading and writing requirements may be impacting students' scores.

Recommendation 6: Routinize Timely Reviews of Local Performance Tasks

As the pool of locally developed tasks expands, it will be important to ensure that the tasks and rubrics are of sufficient quality to be used to generate student scores and annual determinations. Teachers report that their skill level in developing these tasks improves with each year of PACE participation, so it stands to reason that the validity and reliability of students' scores should improve with time. Instituting a system of regular task review will help ensure that happens. A mechanism is in place to evaluate the quality of the locally developed performance tasks and rubrics. Some reviews have been completed at this time (by the New Hampshire Department of Education or by Stanford University), but teachers were frustrated by a lack of feedback from some of these reviews. Review of local tasks would benefit from a regularly scheduled and timely process.

Starting in the 2016-17 school year, districts will be required to submit one major assessment per competency per course, in addition to all local performance tasks in a common task template. At this stage in the evaluation, it is unknown how the assessments/tasks collected during the coming year will be reviewed, what feedback will be available to teachers/schools, or when that feedback will be provided. As this data becomes available, it will be very important to monitor the ways that feedback to teachers/schools is interpreted and used. This process has the potential be very useful and positive for the PACE program, but it also has the potential to introduce unintended consequences.

Recommendations for Interim Goal 3: Performance Assessments Are Successfully Implemented

Recommendation 7: Plan for Future Research on the Impact of PACE on Teaching and Learning

The positive impacts of PACE on teaching and learning should be externally verified. This may be part of a future research agenda when it becomes possible to evaluate the predictive strength of PACE results on college and career performance. In the interim, it may be possible to compare PACE versus non-PACE student performance on Smarter Balanced assessments, college entrance exams, or other measures.

Recommendation 8: Evaluate the Benefit of Time in Program on Outcomes

As the program expands, it may be possible to investigate the benefits of time in the program on instructional practice and student learning. If there is a benefit to spending several years in the PACE program, that may bolster district-level support for the program and promote fidelity of implementation by educators. Teachers described a long period of adjustment and evolution of their teaching and assessment practices. It would not be surprising if there was a direct correlation between years in the program and benefits, both perceived and realized.

Recommendations for Interim Goal 4: Scores Are Accurate and Reliable

Recommendation 9: Consider Systematically Recycling Tasks

After the operational year, common tasks may still be used in place of, or in addition to, locally developed tasks. PACE should consider some method of systematically repeating tasks across years as another check on the consistency of scoring. If tasks were repeated, previously scored "check sets" of student work from the prior year could be included in the current year. Score consistency across years could then be checked in a more systematic way. We recognize that schools have more flexibility in their use of these tasks after their operational year. For example, a task may be administered at a different grade level where the curriculum is better aligned with the task or the task itself may be modified in some ways. Any recycling of tasks must be undertaken with care to ensure such variations do not contaminate the results.

Recommendation 10: Begin Tracking Performance from Year to Year

The PACE system has the potential for variability across years. Comparing performance across years will allow PACE to see where there are large changes in the proportions of students at each achievement level in any district and to investigate potential reasons for those changes. It is important to consider how changes in performance are reported and how they are characterized. Early reports to USED comparing student performance on PACE with performance on Smarter Balanced within and across years[17], as well as the data analyses completed for this evaluation, should be repeated annually. This will allow for continuous monitoring and by investigating anomalous results, PACE may be better able to identify potential threats to reliability and validity. Examples from this report include the lower correlations and reversed convergent/discriminant validity coefficient pattern for grades 7 and 8, as well as larger than typical gains in math for grade 3. Conducting these analyses again next year will help PACE determine if these anomalies are random or if they represent some systematic difference in the way PACE is implemented by grade or subject.
We also recommend that PACE provide guidance for making valid inferences from annual performance information to schools, districts, and, if possible, the media.

End Goal: Students are College and Career Ready

Graduating students who are college and career ready is the ultimate goal of PACE. While we have found considerable evidence supporting the interim goals of PACE, it is still too early to evaluate college and career readiness. Once PACE has matured sufficiently that there are students who both experienced the PACE program and have at least one year of college or career, we recommend that PACE support an ongoing research agenda to investigate claims under this ultimate goal.

[17] See https://www.education.nh.gov/assessment-systems/documents/overview.pdf.

References

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

Becker, D. E., Thacker, A. A., & Sinclair, A. L. (2016a). *Formative evaluation of New Hampshire's Performance Assessment of Competency Education (PACE): June 2016 interim report* (2016 No. 041). Alexandria, VA: Human Resources Research Organization.

Becker, D. E., Thacker, A. A., Sinclair, A. L., Wiley, C. R. H., Woods, A., & Dickinson, E. R. (2016b). *Formative evaluation of New Hampshire's Performance Assessment of Competency Education (PACE): September 2016 interim report* (2016 No. 055). Alexandria, VA: Human Resources Research Organization.

Becker, D. E., Thacker, A. A., Dickinson, E. R., & Sinclair, A. L. (2016c). *Formative evaluation of New Hampshire's Performance Assessment of Competency Education (PACE): Formative Evaluation: Final Report* (2016 No. 085). Alexandria, VA: Human Resources Research Organization.

Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences, 2nd Edition.* Hillsdale: Lawrence Erlbaum.

Dickinson, E. R., & Thacker, A. A. (2009). *Relations among Kentucky's Core Content Test ACT scores and students' self-reported high school grades* (FR-09-32). Alexandria, VA: Human Resources Research Organization.

Evans, C. & Lyons, S. (2016). *NH PACE: Inter-rater reliability analysis report.* Dover, NH: Center for Assessment.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1-73.

Appendix A: Survey Items for Teachers Administering NH PACE Common Performance Tasks

1. Did you/are you administering an NH PACE Common Performance Task(s) in mathematics, science, and/or ELA in grades 3 – 8 or high school any time in 2015-16 or 2016-17?

- o Yes
- o No

Background Items:

2. Please indicate the grade level(s) you currently teach (*Select all that apply)*:

- o Grade 3
- o Grade 4
- o Grade 5
- o Grade 6
- o Grade 7
- o Grade 8
- o Grade 9
- o Grade 10
- o Grade 11
- o Grade 12

3. Please indicate the content area(s) for which you are responsible for administering NH PACE Common Performance Task(s) in the 2016–17 school year (*Select all that apply)*:

- o ELA
- o mathematics
- o science

4. Please indicate how long you personally have been a part of the PACE pilot program:

- o Since the 2014-15 school year
- o Since the 2015-16 school year
- o Since the 2016–17 school year

5. Have you switched schools or districts since the 2014-15 school year (*Select all that apply)*?

- o No
- o Yes, I switched schools and/or districts in 2014-15.
- o Yes, I switched schools and/or districts in 2015-16.
- o Yes, I switched schools and/or districts in 2016-17.

Instructions: Please respond to the survey items based on your current impression of the PACE Pilot Program.

Please rate your level of agreement with the following statements:

SD = Strongly disagree;
D = Disagree;
N = Neutral;
A = Agree;
SA = Strongly agree
DK = Don't Know

6.	SD	D	N	A	SA	DK
a. My school's administration (e.g., principal, assistant principal, curriculum director) is supportive of the PACE initiative.						
b. My school's administration provides me with the resources and supports that I need to effectively implement the NH PACE Common Performance Tasks.						
c. The information I receive from my school's administration about the PACE requirements and PACE updates is communicated to me in an effective and timely manner.						
d. The teachers at my school effectively collaborate with one another on topics relevant to the implementation of the PACE pilot.						
e. The time and effort required by the PACE initiative are worth the benefits that I have experienced and/or seen.						

7. Think about the teachers at your school who are familiar with the PACE initiative; this includes teachers who administer NH PACE common performance tasks AND teachers who do not administer common performance tasks, but who are familiar with the PACE initiative via attendance at faculty meetings, informal discussions with teachers administering PACE tasks, etc. Please select the statement that most closely reflects your perception of their opinion about PACE.

- o They all have a favorable opinion of the PACE initiative
- o Most have a favorable opinion of the PACE initiative
- o Some have a favorable opinion of the PACE initiative
- o Few have a favorable opinion of the PACE initiative
- o None have a favorable opinion of the PACE initiative
- o Don't know

8. Have you participated in any cross-district collaborations on NH PACE Common Performance Tasks (i.e., task development sessions and/or calibration sessions)?

- o Yes
- o No *[Note: If "No," skip to Question 15]*

9. Aside from face-to-face meetings, identify the resources/methods that have been most effective for facilitating useful cross-district collaborations. Enter a '1' for the most effective resource/method; enter a '2' for the second most effective resource/method, enter a '3' for the third most effective resource/method, and so forth. Enter a '9' for any resources you have not used.

- o ___ LibGuides
- o ___ Emails
- o ___ Teleconferences
- o ___ Google docs
- o ____ Other: _____

10. Please indicate the extent to which collaborations with teachers from **other** districts have been useful for developing NH PACE Common Performance Tasks.

- o Not useful
- o Slightly useful
- o Somewhat useful
- o Very useful
- o Extremely useful
- o Not applicable (I have not participated in cross-district task development sessions)

11. Please indicate the extent to which collaborations with teachers from **other** districts have been useful for calibrating the scoring of student work on the NH PACE Common Performance Tasks.

- o Not useful
- o Slightly useful
- o Somewhat useful
- o Very useful
- o Extremely useful
- o Not applicable (I have not participated in cross-district calibration sessions)

12. Have you been selected by NH DOE to serve in an official capacity as a "PACE Content Lead" to facilitate and support your colleagues in the development of pilot and operational NH PACE Common Performance Tasks?

- o Yes
- o No

13. Do you currently serve as a "Teacher Representative" for NH PACE Common Performance Tasks? That is, are you participating in cross-district task development sessions and communicating progress on task development to other teachers in your district?

- o Yes
- o No

14. Please indicate the extent to which the Content Leads have provided useful guidance and support in . . .	Not Useful	Slightly Useful	Somewhat Useful	Very Useful	Extremely Useful	NA
a. facilitating task development sessions						
b. organizing the materials needed for task development						
c. communicating information about requirements for task development						
d. answering questions about scaffolding						
e. answering questions about the rubric						

Please rate your level of agreement with the following statements:

SD = Strongly disagree;
D = Disagree;
N = Neutral;
A = Agree;
SA = Strongly agree;
DK = Don't know;
NA = Not applicable

15.	**SD**	**D**	**N**	**A**	**SA**	**DK**	**NA**
a. The NH PACE Common Performance Tasks are authentic measures of my students' achievement.							
b. The locally developed performance tasks (i.e., "non-common tasks") are authentic measures of my students' achievement.							
c. I have been able to apply what I've learned from the NH PACE Common Performance Tasks to developing higher quality local performance tasks (i.e., "non-common tasks").							

16. Number of local performance tasks (i.e., "non-common tasks") that I plan to administer this year for SCIENCE:

- o Zero
- o 1 – 5
- o 6 – 10
- o More than 10

17. Total number of local performance tasks (i.e., "non-common tasks") that I plan to administer this year for MATH:

- o Zero
- o 1 – 5
- o 6 – 10
- o More than 10

18. Total number of local performance tasks (i.e., “non-common tasks”) that I plan to administer this year for ELA:

- o Zero
- o 1 – 5
- o 6 – 10
- o More than 10

19. Total number of local performance tasks (i.e., “non-common tasks”) that I plan to administer this year for OTHER content areas (e.g., Social Studies):

- o Zero
- o 1 – 5
- o 6 – 10
- o More than 10

Please rate your level of agreement with the following statements:

SD = Strongly disagree;
D = Disagree;
N = Neutral;
A = Agree;
SA = Strongly agree;
DK = Don’t know;
NA = Not applicable*;*

20.	**SD**	**D**	**N**	**A**	**SA**	**DK**	**NA**
a. I have received adequate training and preparation to effectively administer the NH PACE Common Performance Tasks.							
b. I understand, based on the scaffolding rules, the amount of scaffolding I can employ with the NH PACE Common Performance Tasks.							
c. The NH PACE Implementation Guidelines Manual has been a useful resource in helping me understand how to effectively administer the NH PACE Common Performance Tasks.							
d. The NH PACE Accommodations Guidelines have been useful resources in helping me understand the appropriate accommodations for the NH PACE Common Performance Tasks.							
e. From my perspective, different teachers in different classrooms are administering the same NH PACE Common Performance Tasks in a consistent manner.							

According to Sturgis and Patrick (2011), competency-based education is characterized by the following:

- Competencies include explicit, measurable, transferable learning objectives that empower students.
- Assessment is meaningful and a positive learning experience for students.
- Students receive rapid, differentiated support based on their individual learning needs.
- Learning outcomes emphasize competencies that include the application and creation of knowledge.
- The process of reaching learning outcomes encourages students to develop skills and dispositions important for success in college, careers and citizenship.

21. Based on the above characterization of competency-based education, indicate the extent to which competency-based education is integrated into your instruction.

- o To no extent
- o To a slight extent
- o To some extent
- o To a great extent
- o To a very great extent
- o Not applicable

22. I use performance tasks for **instructional purposes**:

- o Never
- o A couple times a month or less
- o Approximately once each week
- o 2 – 3 times each week
- o Nearly every day

23. I use performance tasks for **assessment purposes:**

- o Never
- o A couple times a month or less
- o Approximately once each week
- o 2 – 3 times each week
- o Nearly every day

Read each of the following statements. Then, rate your level of agreement with each statement for (a) your classroom (i.e., classroom impact) and (b) your school (i.e., schoolwide impact).

SD = Strongly disagree;
D = Disagree;
N = Neutral;
A = Agree;
SA = Strongly agree;
DK = Don't know

24.	1_In My Classroom						2_In My School					
Implementing performance tasks has had a positive impact on . . .	**SD**	**D**	**N**	**A**	**SA**	**DK**	**SD**	**D**	**N**	**A**	**SA**	**DK**
a. instructional practice, such that instruction occurs at a higher depth of knowledge.												
b. student engagement while completing performance tasks.												
c. student engagement in learning overall.												

Please rate your level of agreement with the following statements:

SD = Strongly disagree;
D = Disagree;
N = Neutral;
A = Agree;
SA = Strongly agree
DK = Don't Know

25.	SD	D	N	A	SA	DK
a. The scoring rubrics for the NH PACE Common Performance Tasks are sufficiently clear and detailed to ensure that separate scorers scoring the same student work arrive at the same score.						
b. The scoring resources available on the LibGuide effectively explain how to score the student work on the NH PACE Common Performance Tasks.						
c. The **within-district** double scoring is effective in ensuring the reliability of scores on the NH PACE Common Performance Tasks.						

26. The NH PACE Common Performance Tasks are more accessible to a greater range of student learning needs (e.g., students with disabilities, English language learners) than traditional standardized tests.

- o Strongly disagree
- o Disagree
- o Neutral
- o Agree5
- o Strong agree
- o Don't know

27. Please use the space below to provide any additional information you would like to share about PACE.

Appendix B: Survey Items for Teachers Not Administering NH PACE Common Performance Tasks

Background Items:

1. Are you currently teaching students in a particular grade level(s) or content area (e.g., kindergarten teacher, 6th grade social studies, elementary school music teacher)?

- o Yes
- o No [18]

2. Please indicate the grade level(s) you currently teach (*Select all that apply)*:

- o Kindergarten
- o Grade 1
- o Grade 2
- o Grade 3
- o Grade 4
- o Grade 5
- o Grade 6
- o Grade 7
- o Grade 8
- o Grade 9
- o Grade 10
- o Grade 11
- o Grade 12

3. Please indicate the content area(s) for which you currently teach (*Select all that apply)*:

- o ELA
- o mathematics
- o science
- o Art
- o Music/Fine Arts
- o Social Studies/History
- o World Language
- o Special Education
- o Vocational Studies
- o Physical Education/Health
- o Library
- o Drama
- o Computers/Technology
- o __ Other _____

[18] If the respondent selects "No," then he/she is exited from the survey and taken to a message that says, "The remainder of the survey items are specific to teachers teaching a particular content area and/or grade level. Consequently, no additional input is needed from you at this time. Thank-you for your time!"

This additional item was added to "weed out" any email addresses that are not teachers (e.g., guidance counselors, speech pathologists, etc.).

4. Have you switched schools or districts since the 2014-15 school year (*Select all that apply)*?

- o No
- o Yes, I switched schools and/or districts in 2014-15.
- o Yes, I switched schools and/or districts in 2015-16.
- o Yes, I switched schools and/or districts in 2016-17.

5. Please rate your level of familiarity with the NH Performance Assessment of Competency Education (PACE) pilot program at your school.

- o Unfamiliar
- o Somewhat unfamiliar
- o Somewhat familiar
- o Very familiar
- o Extremely familiar

6. Do you develop and/or administer local performance tasks (i.e., "non-common tasks") in the content area(s)/grade(s) you teach?

- o Yes
- o No

7. Please indicate the approximate number of local performance tasks (i.e., "non-common tasks") that you plan to administer this academic year.

- o Zero
- o 1 – 5
- o 6 – 10
- o More than 10

Instructions: Please respond to the survey items based on your current impression of the PACE Pilot Program.

8. My school's administration (e.g., principal, assistant principal, curriculum director) is supportive of the PACE initiative.

- o Strongly disagree
- o Disagree
- o Neutral
- o Agree
- o Strongly agree
- o Don't know

9. Think about the teachers at your school who are familiar with the PACE initiative; this includes teachers who administer NH PACE common performance tasks AND teachers who do not administer common performance tasks, but who are familiar with the PACE initiative via attendance at faculty meetings, informal discussions with teachers administering PACE tasks, etc. Please select the statement that most closely reflects your perception of their opinion about PACE.

- o They all have a favorable opinion of the PACE initiative
- o Most have a favorable opinion of the PACE initiative
- o Some have a favorable opinion of the PACE initiative
- o Few have a favorable opinion of the PACE initiative
- o None have a favorable opinion of the PACE initiative
- o Don't know

According to Sturgis and Patrick (2011), competency-based education is characterized by the following:

- Competencies include explicit, measurable, transferable learning objectives that empower students.
- Assessment is meaningful and a positive learning experience for students.
- Students receive rapid, differentiated support based on their individual learning needs.
- Learning outcomes emphasize competencies that include the application and creation of knowledge.
- The process of reaching learning outcomes encourages students to develop skills and dispositions important for success in college, careers and citizenship.

10. Based on the above characterization of competency-based education, indicate the extent to which competency-based education is integrated into your instruction.

- o To no extent
- o To a slight extent
- o To some extent
- o To a great extent
- o To a very great extent
- o Not applicable

11. I use performance tasks for **instructional purposes**:

- o Never
- o A couple times a month or less
- o Approximately once each week
- o 2 – 3 times each week
- o Nearly every day

12. I use performance tasks for **assessment purposes:**

- o Never
- o A couple times a month or less
- o Approximately once each week
- o 2 – 3 times each week
- o Nearly every day

Read each of the following statements. Then, rate your level of agreement with each statement for (a) your classroom (i.e., classroom impact) and (b) your school (i.e., schoolwide impact).

SD = Strongly disagree;
D = Disagree;
N = Neutral;
A = Agree;
SA = Strongly agree;
DK = Don't know

13.	1_In My Classroom						2_In My School					
Implementing performance tasks has had a positive impact on . . .	SD	D	N	A	SA	DK	SD	D	N	A	SA	DK
a. instructional practice, such that instruction occurs at a higher depth of knowledge.												
b. student engagement while completing performance tasks.												
c. student engagement in learning overall.												

Appendix C: PACE Teacher Survey Response Frequencies

Table C1. PACE Teachers Survey: Q2. Please indicate the grade level(s) you currently teach. Select all that apply.

	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 9		Grade 10		Grade 11		Grade 12	
District	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Concord	8	11.9	10	14.9	14	20.9	12	17.9	8	11.9	3	4.5	13	19.4	17	25.4	11	16.4	11	16.4
Epping	4	17.4	6	26.1	7	30.4	3	13.0	3	13.0	2	8.7	4	17.4	4	17.4	2	8.7	2	8.7
Rochester	12	20.7	15	25.9	8	13.8	5	8.6	4	6.9	3	5.2	17	29.3	18	31.0	12	20.7	10	17.2
Sanborn	6	14.6	6	14.6	8	19.5	3	7.3	6	14.6	2	4.9	10	24.4	12	29.3	11	26.8	12	29.3
Souhegan	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	9	42.9	13	61.9	11	52.4	10	47.6
White Mountains	7	23.3	7	23.3	8	26.7	8	26.7	6	20.0	6	20.0	6	20.0	8	26.7	7	23.3	7	23.3
Small Districts Combined[a]	6	20.7	6	20.7	7	24.1	6	20.7	8	27.6	10	35.4	5	17.2	6	20.7	4	13.8	2	6.9
All Districts	43	16.0	50	18.6	52	19.3	37	13.8	35	13.0	26	9.7	64	23.8	78	29.0	58	21.6	54	20.1

[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C2. PACE Teachers Survey: Q3. Please indicate the content area(s) for which you are responsible for administering NH PACE Common Performance Task(s) in the 2016–17 school year. Select all that apply.

	ELA		Mathematics		Science	
District	n	%	n	%	n	%
Concord	30	44.8	37	55.2	22	32.8
Epping	11	47.8	12	52.2	7	30.4
Rochester	33	56.9	26	44.8	21	36.2
Sanborn	22	53.7	20	48.8	11	26.8
Souhegan	6	28.6	8	38.1	8	38.1
White Mountains	18	60.0	15	50.0	10	33.3
Small Districts Combined[a]	15	51.7	19	65.5	11	37.9
All Districts	135	50.2	137	50.9	90	33.5

Note. ELA= English Language Arts

[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C3. PACE Teachers Survey: Q4. Please indicate how long you personally have been a part of the PACE pilot program.

	Since the 2014-15 School Year		Since the 2015-16 School Year		Since the 2016–17 School Year	
District	n	%	n	%	n	%
Concord	15	22.4	46	68.7	6	9.0
Epping	19	82.6	2	8.7	2	8.7
Rochester	42	72.4	8	13.8	8	13.8
Sanborn	29	70.7	6	14.6	6	14.6
Souhegan	20	95.2	1	4.8	0	0.0
White Mountains	1	3.3	2	6.7	27	90.0
Small Districts Combined[a]	7	24.1	18	62.1	4	13.8
All Districts	133	49.4	83	30.9	53	19.7

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C4. PACE Teachers Survey: Q5. Have you switched schools or districts since the 2014—15 school year? Select all that apply.

	No		Yes, I switched schools and/or districts in 2014-15		Yes, I switched schools and/or districts in 2015-16		Yes, I switched schools and/or districts in 2016-17	
District	n	%	n	%	n	%	n	%
Concord	66	98.5	0	0.0	1	1.5	0	0.0
Epping	20	87.0	0	0.0	2	8.7	1	4.3
Rochester	50	86.2	3	5.2	5	8.6	3	5.2
Sanborn	36	87.8	1	2.4	3	7.3	3	7.3
Souhegan	21	100.0	0	0.0	0	0.0	0	0.0
White Mountains	23	76.0	2	6.7	0	0.0	5	16.7
Small Districts Combined[a]	25	86.2	0	0.0	2	6.9	2	6.9
All Districts	241	89.6	6	2.2	13	4.8	14	5.2

[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C5. PACE Teachers Survey: Q6a. Please rate your level of agreement with the following statement: My school's administration (e.g., principal, assistant principal, curriculum director) is supportive of the PACE initiative.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	1	1.5	2	3.0	12	17.9	48	71.6	4	6.0
Epping	0	0.0	0	0.0	3	13.0	9	39.1	10	43.5	1	4.3
Rochester	0	0.0	0	0.0	5	8.6	17	29.3	32	55.2	4	6.9
Sanborn	0	0.0	1	2.4	1	2.4	7	17.1	32	78.0	0	0.0
Souhegan	0	0.0	0	0.0	0	0.0	0	0.0	21	100.0	0	0.0
White Mountains	0	0.0	0	0.0	3	10.0	8	26.7	18	60.0	1	3.3
Small Districts Combined[a]	0	0.0	0	0.0	2	6.9	9	31.0	18	62.1	0	0.0
All Districts	0	0.0	2	0.7	16	5.9	62	23.0	179	66.5	10	3.7

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C6. PACE Teachers Survey: Q6b. Please rate your level of agreement with the following statement: My school's administration provides me with the resources and supports that I need to effectively implement the NH PACE Common Performance Tasks.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	2	3.0	10	14.9	27	40.3	28	41.8	0	0.0
Epping	0	0.0	1	4.3	5	21.7	9	39.1	8	34.8	0	0.0
Rochester	1	1.7	6	10.3	5	8.6	21	36.2	23	39.7	2	3.4
Sanborn	0	0.0	0	0.0	4	9.8	13	31.7	23	56.1	1	2.4
Souhegan	0	0.0	0	0.0	0	0.0	7	33.3	14	66.7	0	0.0
White Mountains	0	0.0	3	10.0	3	10.0	17	56.7	7	23.3	0	0.0
Small Districts Combined[a]	0	0.0	3	10.3	4	13.8	17	58.6	5	17.2	0	0.0
All Districts	1	0.4	15	5.6	31	11.5	111	41.3	108	40.1	3	1.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C7. PACE Teachers Survey: Q6c. Please rate your level of agreement with the following statement: The information I receive from my school's administration about the PACE requirements and PACE updates is communicated to me in an effective and timely manner.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	1	1.5	1	1.5	7	10.4	35	52.2	23	34.3	0	0.0
Epping	1	4.3	7	30.4	4	17.4	7	30.4	4	17.4	0	0.0
Rochester	3	5.2	10	17.2	10	17.2	19	32.8	14	24.1	2	3.4
Sanborn	2	4.9	4	9.8	10	24.4	14	34.1	10	24.4	1	2.4
Souhegan	0	0.0	1	4.8	2	9.5	8	38.1	10	47.6	0	0.0
White Mountains	0	0.0	4	13.3	3	10.0	14	46.7	9	30.0	0	0.0
Small Districts Combined[a]	1	3.4	2	6.9	2	6.9	16	55.2	8	27.6	0	0.0
All Districts	8	3.0	29	10.8	38	14.1	113	42.0	78	29.0	3	1.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C8. PACE Teachers Survey: Q6d. Please rate your level of agreement with the following statement: The teachers at my school effectively collaborate with one another on topics relevant to the implementation of the PACE pilot.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	1	1.5	5	7.5	27	40.3	34	50.7	0	0.0
Epping	2	8.7	2	8.7	5	21.7	8	34.8	6	26.1	0	0.0
Rochester	0	0.0	3	5.2	3	5.2	21	36.2	29	50.0	2	3.4
Sanborn	0	0.0	1	2.4	5	12.2	17	41.5	17	41.5	1	2.4
Souhegan	0	0.0	0	0.0	1	4.8	1	4.8	19	90.5	0	0.0
White Mountains	0	0.0	3	10.0	4	13.3	14	46.7	9	30.0	0	0.0
Small Districts Combined[a]	1	3.4	5	17.2	6	20.7	11	37.9	6	20.7	0	0.0
All Districts	3	1.1	15	5.6	29	10.8	99	36.8	120	44.6	3	1.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C9. PACE Teachers Survey: Q6e. Please rate your level of agreement with the following statement: The time and effort required by the PACE initiative are worth the benefits that I have experienced and/or seen.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	3	4.5	7	10.4	18	26.9	20	29.9	15	22.4	4	6.0
Epping	4	17.4	5	21.7	5	21.7	7	30.4	1	4.3	1	4.3
Rochester	8	13.8	9	15.5	16	27.6	13	22.4	8	13.8	4	6.9
Sanborn	3	7.3	3	7.3	5	12.2	14	34.1	15	36.6	1	2.4
Souhegan	0	0.0	0	0.0	3	14.3	9	42.9	8	38.1	1	4.8
White Mountains	7	23.3	3	10.0	9	30.0	2	6.7	1	3.3	8	26.7
Small Districts Combined[a]	4	13.8	9	31.0	5	17.2	9	31.0	1	3.4	1	3.4
All Districts	29	10.8	36	13.4	61	22.7	74	27.5	49	18.2	20	7.4

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C10. PACE Teachers Survey: Q7. Think about the teachers at your school who are familiar with the PACE initiative; this includes teachers who administer NH PACE common performance tasks AND teachers who do not administer common performance tasks, but who are familiar with the PACE initiative via attendance at faculty meetings, informal discussions with teachers administering PACE tasks, etc. Please select the statement that most closely reflects your perception of their opinion about PACE.

	They all have a favorable opinion of the PACE initiative.		Most have a favorable opinion of the PACE initiative		Some have a favorable opinion of the PACE initiative		Few have a favorable opinion of the PACE initiative		None have a favorable opinion of the PACE initiative		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	3	4.5	32	47.8	22	32.8	3	4.5	0	0.0	7	10.4
Epping	0	0.0	8	34.8	7	30.4	7	30.4	1	4.3	0	0.0
Rochester	4	6.9	13	22.4	28	48.3	9	15.5	1	1.7	3	5.2
Sanborn	3	7.3	20	48.8	10	24.4	7	17.1	0	0.0	1	2.4
Souhegan	6	28.6	11	52.4	2	9.5	1	4.8	0	0.0	1	4.8
White Mountains	0	0.0	3	10.0	13	43.3	12	40.0	1	3.3	1	3.3
Small Districts Combined[a]	0	0.0	8	27.6	12	41.4	6	20.7	1	3.4	2	6.9
All Districts	16	5.9	95	35.3	94	34.9	45	16.7	4	1.5	15	5.6

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C11. PACE Teachers Survey: Q8. Have you participated in any cross-district collaborations on NH PACE Common Performance Tasks (i.e., task development sessions and/or calibration sessions)?

	Yes		No	
District	n	%	n	%
Concord	42	62.7	25	37.3
Epping	20	87.0	3	13.0
Rochester	31	53.4	27	46.6
Sanborn	35	85.4	6	14.6
Souhegan	16	76.2	5	23.8
White Mountains	21	70.0	9	30.0
Small Districts Combined[a]	17	58.6	12	41.4
All Districts	182	67.7	87	32.3

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C12. PACE Teachers Survey: Q9a. Aside from face-to-face meetings, identify the resources/methods that have been most effective for facilitating useful cross-district collaborations: LibGuides

	Most Effective Resource/ Method		Second Most Effective Resource/ Method		Third Most Effective Resource/ Method		Fourth Most Effective Resource/ Method		Have Not Used This Resource/ Method	
District	n	%	n	%	n	%	n	%	n	%
Concord	9	21.4	10	23.8	8	19.0	1	2.4	10	23.8
Epping	8	40.0	5	25.0	5	25.0	0	0.0	2	10.0
Rochester	8	25.0	4	12.5	6	18.8	2	6.3	12	37.5
Sanborn	12	34.3	11	31.4	8	22.9	2	5.7	2	5.7
Souhegan	3	18.8	7	43.8	3	18.8	1	6.3	2	12.5
White Mountains	12	57.1	2	9.5	4	19.0	1	4.8	2	9.5
Small Districts Combined[a]	10	55.6	5	27.8	1	5.6	2	11.1	0	0.0
All Districts	62	33.7	44	23.9	35	19.0	9	4.9	30	16.3

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C13. PACE Teachers Survey: Q9b. Aside from face-to-face meetings, identify the resources/methods that have been most effective for facilitating useful cross-district collaborations: Emails

	Most Effective Resource/ Method		Second Most Effective Resource/ Method		Third Most Effective Resource/ Method		Fourth Most Effective Resource/ Method		Have Not Used This Resource/ Method	
District	n	%	n	%	n	%	n	%	n	%
Concord	14	33.3	16	38.1	8	19.0	1	2.4	2	4.8
Epping	5	26.3	8	42.1	4	21.1	0	0.0	2	10.5
Rochester	13	40.6	9	28.1	7	21.9	1	3.1	2	6.3
Sanborn	8	22.9	15	42.9	11	31.4	1	2.9	0	0.0
Souhegan	3	18.8	5	31.3	7	43.8	1	6.3	0	0.0
White Mountains	4	19.0	11	52.4	3	14.3	3	14.3	0	0.0
Small Districts Combined[a]	5	27.8	5	27.8	8	44.4	0	0.0	0	0.0
All Districts	52	28.4	69	37.7	48	26.2	7	3.8	6	3.3

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C14. PACE Teachers Survey: Q9c. Aside from face-to-face meetings, identify the resources/methods that have been most effective for facilitating useful cross-district collaborations: Teleconferences

	Most Effective Resource/ Method		Second Most Effective Resource/ Method		Third Most Effective Resource/ Method		Fourth Most Effective Resource/ Method		Have Not Used This Resource/ Method	
District	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	0	0.0	1	2.4	12	28.6	29	69.0
Epping	0	0.0	0	0.0	0	0.0	3	15.8	16	84.2
Rochester	0	0.0	0	0.0	4	12.5	1	3.1	27	84.4
Sanborn	0	0.0	0	0.0	1	2.9	5	14.7	27	79.4
Souhegan	0	0.0	1	6.3	2	12.5	4	25.0	9	56.3
White Mountains	0	0.0	0	0.0	1	4.8	1	4.8	17	81.0
Small Districts Combined[a]	1	5.6	0	0.0	1	5.6	3	16.7	12	66.7
All Districts	1	0.5	1	0.5	10	5.5	29	15.9	137	75.3

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C15. PACE Teachers Survey: Q9d. Aside from face-to-face meetings, identify the resources/methods that have been most effective for facilitating useful cross-district collaborations: Google Docs

	Most Effective Resource/ Method		Second Most Effective Resource/ Method		Third Most Effective Resource/ Method		Fourth Most Effective Resource/ Method		Have Not Used This Resource/ Method	
District	n	%	n	%	n	%	n	%	n	%
Concord	12	28.6	9	21.4	12	28.6	0	0.0	8	19.0
Epping	5	26.3	5	26.3	6	31.6	1	5.3	2	10.5
Rochester	7	21.9	16	50.0	5	15.6	1	3.1	3	9.4
Sanborn	15	42.9	8	22.9	11	31.4	0	0.0	1	2.9
Souhegan	10	62.5	3	18.8	1	6.3	0	0.0	2	12.5
White Mountains	3	14.3	5	23.8	10	47.6	0	0.0	2	9.5
Small Districts Combined[a]	1	5.6	8	44.4	4	22.2	1	5.6	4	22.2
All Districts	53	29.0	54	29.5	49	26.8	3	1.6	22	12.0

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C16. PACE Teachers Survey: Q10. Please indicate the extent to which collaborations with teachers from other districts have been useful for developing NH PACE Common Performance Tasks.

	Not Useful		Slightly Useful		Somewhat Useful		Very Useful		Extremely Useful		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	1	2.4	3	7.1	10	23.8	15	35.7	10	23.8	3	7.1
Epping	1	5.0	0	0.0	6	30.0	7	35.0	6	30.0	0	0.0
Rochester	1	3.2	0	0.0	4	12.9	10	32.3	10	32.3	6	19.4
Sanborn	1	2.9	2	5.7	8	22.9	12	34.3	12	34.3	0	0.0
Souhegan	0	0.0	0	0.0	1	6.3	6	37.5	9	56.3	0	0.0
White Mountains	2	9.5	3	14.3	7	33.3	6	28.6	2	9.5	1	4.8
Small Districts Combined[a]	0	0.0	2	11.8	5	29.4	5	29.4	4	23.5	1	5.9
All Districts	6	3.3	10	5.5	41	22.5	61	33.5	53	29.1	11	6.0

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C17. PACE Teachers Survey: Q11. Please indicate the extent to which collaborations with teachers from other districts have been useful for calibrating the scoring of student work on the NH PACE Common Performance Tasks.

	Not Useful		Slightly Useful		Somewhat Useful		Very Useful		Extremely Useful		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	3	7.1	3	7.1	8	19.0	8	19.0	7	16.7	13	31.0
Epping	0	0.0	1	5.0	6	30.0	9	45.0	3	15.0	1	5.0
Rochester	1	3.2	0	0.0	1	3.2	11	35.5	8	25.8	10	32.3
Sanborn	1	2.9	0	0.0	8	22.9	16	45.7	7	20.0	3	8.6
Souhegan	0	0.0	1	6.3	0	0.0	5	31.3	5	31.3	5	31.3
White Mountains	3	14.3	2	9.5	2	9.5	3	14.3	1	4.8	10	47.6
Small Districts Combined[a]	0	0.0	3	17.6	6	35.3	5	29.4	2	11.8	1	5.9
All Districts	8	4.4	10	5.5	31	17.0	57	31.3	33	18.1	43	23.6

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C18. PACE Teachers Survey: Q12. Have you been selected by NH DOE to serve in an official capacity as a "PACE Content Lead" to facilitate and support your colleagues in the development of pilot and operational NH PACE Common Performance Tasks?

	Yes		No	
District	n	%	n	%
Concord	5	11.9	37	88.1
Epping	1	5.0	19	95.0
Rochester	4	12.9	27	87.1
Sanborn	8	22.9	27	77.1
Souhegan	6	37.5	10	62.5
White Mountains	2	9.5	19	90.5
Small Districts Combined[a]	3	17.6	14	82.4
All Districts	29	15.9	153	84.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C19. PACE Teachers Survey: Q13. Do you currently serve as a "Teacher Representative" for NH PACE Common Performance Tasks? That is, are you participating in cross-district task development sessions and communicating progress on task development to other teachers in your district?

	Yes		No	
District	n	%	n	%
Concord	28	66.7	14	33.3
Epping	16	80.0	4	20.0
Rochester	15	48.4	16	51.6
Sanborn	20	57.1	15	42.9
Souhegan	12	75.0	4	25.0
White Mountains	13	61.9	8	38.1
Small Districts Combined[a]	12	70.6	5	29.4
All Districts	116	63.7	66	36.3

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C20. PACE Teachers Survey: Q14a. Please indicate the extent to which the Content Leads have provided useful guidance and support in facilitating task development sessions.

	Not Useful		Slightly Useful		Somewhat Useful		Very Useful		Extremely Useful		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	2	4.8	1	2.4	6	14.3	17	40.5	10	23.8	6	14.3
Epping	1	5.0	5	25.5	5	25.5	6	30.0	3	15.0	0	0.0
Rochester	2	6.5	1	3.2	8	25.8	7	22.6	7	22.6	6	19.4
Sanborn	1	2.9	2	5.7	10	28.6	8	22.9	11	31.4	3	8.6
Souhegan	0	0.0	0	0.0	0	0.0	7	43.8	8	50.0	1	6.3
White Mountains	0	0.0	1	4.8	3	14.3	10	47.6	5	23.8	2	9.5
Small Districts Combined[a]	1	5.9	2	11.8	6	35.3	3	17.6	4	23.5	1	5.9
All Districts	7	3.8	12	6.6	38	20.9	58	31.9	48	26.4	19	10.4

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C21. PACE Teachers Survey: Q14b. Please indicate the extent to which the Content Leads have provided useful guidance and support in organizing the materials needed for task development.

	Not Useful		Slightly Useful		Somewhat Useful		Very Useful		Extremely Useful		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	2	4.8	0	0.0	9	21.4	17	40.5	10	23.8	4	9.5
Epping	3	15.0	4	20.0	4	20.0	5	25.5	4	20.0	0	0.0
Rochester	2	6.5	0	0.0	10	32.3	7	22.6	7	22.6	5	16.1
Sanborn	0	0.0	2	5.7	12	34.3	9	25.7	10	28.6	2	5.7
Souhegan	0	0.0	0	0.0	0	0.0	7	43.8	8	50.0	1	6.3
White Mountains	0	0.0	2	9.5	5	23.8	10	47.6	3	14.3	1	4.8
Small Districts Combined[a]	1	5.9	1	5.9	6	35.3	5	29.4	3	17.6	1	5.9
All Districts	8	4.4	9	4.9	46	25.3	60	33.0	45	24.7	14	7.7

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C22. PACE Teachers Survey: Q14c. Please indicate the extent to which the Content Leads have provided useful guidance and support in communicating information about requirements for task development.

	Not Useful		Slightly Useful		Somewhat Useful		Very Useful		Extremely Useful		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	2	4.8	1	2.4	7	16.7	17	40.5	11	26.2	4	9.5
Epping	3	15.0	1	5.0	7	35.0	4	20.0	5	25.0	0	0.0
Rochester	1	3.2	0	0.0	10	32.3	10	32.3	5	16.1	5	16.1
Sanborn	1	2.9	2	5.7	10	28.6	10	28.6	10	28.6	2	5.7
Souhegan	0	0.0	0	0.0	0	0.0	7	43.8	8	50.0	1	6.3
White Mountains	0	0.0	2	9.5	7	33.3	8	38.1	3	14.3	1	4.8
Small Districts Combined[a]	1	5.9	2	11.8	5	29.4	5	29.4	3	17.6	1	5.9
All Districts	8	4.4	8	4.4	46	25.3	61	33.5	45	24.7	14	7.7

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C23. PACE Teachers Survey: Q14d. Please indicate the extent to which the Content Leads have provided useful guidance and support in answering questions about scaffolding.

	Not Useful		Slightly Useful		Somewhat Useful		Very Useful		Extremely Useful		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	3	7.1	3	7.1	10	23.8	18	42.9	3	7.1	5	11.9
Epping	3	15.0	4	20.0	5	25.0	5	25.0	1	5.0	2	10.0
Rochester	1	3.2	3	9.7	13	41.9	6	19.4	4	12.9	4	12.9
Sanborn	2	5.7	2	5.7	11	31.4	8	22.9	10	28.6	2	5.7
Souhegan	0	0.0	0	0.0	2	12.5	8	50.0	5	31.3	1	6.3
White Mountains	1	4.8	2	9.5	5	23.8	5	23.8	4	19.0	4	19.0
Small Districts Combined[a]	1	5.9	3	17.6	5	29.4	5	29.4	2	11.8	1	5.9
All Districts	11	6.0	17	9.3	51	28.0	55	30.2	29	15.9	19	10.4

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C24. PACE Teachers Survey: Q14e. Please indicate the extent to which the Content Leads have provided useful guidance and support in answering questions about the rubric.

	Not Useful		Slightly Useful		Somewhat Useful		Very Useful		Extremely Useful		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	3	7.1	3	7.1	5	11.9	20	47.6	6	14.3	5	11.9
Epping	4	20.0	3	15.0	5	25.0	7	35.0	0	0.0	1	5.0
Rochester	1	3.2	1	3.2	15	48.4	7	22.6	3	9.7	4	12.9
Sanborn	3	8.6	0	0.0	9	25.7	9	25.7	11	31.4	3	8.6
Souhegan	0	0.0	0	0.0	2	12.5	8	50.0	5	31.3	1	6.3
White Mountains	0	0.0	1	4.8	7	33.3	7	33.3	3	14.3	3	14.3
Small Districts Combined[a]	1	5.9	3	17.6	4	23.5	5	29.4	3	17.6	1	5.9
All Districts	12	6.6	11	6.0	47	25.8	63	34.6	31	17.0	18	9.9

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C25. PACE Teachers Survey: Q15a. Please rate your level of agreement with the following statement: The NH PACE Common Performance Tasks are authentic measures of my students' achievement.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	7	10.4	10	14.9	6	9.0	33	49.3	10	14.9	1	1.5
Epping	1	4.3	3	13.0	5	21.7	10	43.5	3	13.0	1	4.3
Rochester	8	13.8	4	6.9	15	25.9	24	41.4	5	8.6	2	3.4
Sanborn	1	2.6	2	5.1	3	7.7	17	43.6	16	41.0	0	0.0
Souhegan	0	0.0	0	0.0	2	9.5	8	38.1	10	47.6	1	4.8
White Mountains	2	6.9	3	10.3	7	24.1	10	34.5	2	6.9	5	17.2
Small Districts Combined[a]	2	7.1	3	10.7	3	10.7	13	46.4	6	21.4	1	3.6
All Districts	21	7.9	25	9.4	41	15.5	115	43.4	52	19.6	11	4.2

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C26. PACE Teachers Survey: Q15b. Please rate your level of agreement with the following statement: The locally developed performance tasks (i.e., "non-common tasks") are authentic measures of my students' achievement.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	5	7.5	3	4.5	12	17.9	33	49.3	12	17.9	2	3.0
Epping	1	4.3	1	4.3	3	13.0	13	56.5	4	17.4	1	4.3
Rochester	4	6.9	4	6.9	14	24.1	25	43.1	7	12.1	4	6.9
Sanborn	0	0.0	2	5.1	5	12.8	13	33.3	19	48.7	0	0.0
Souhegan	0	0.0	0	0.0	0	0.0	5	23.8	16	76.2	0	0.0
White Mountains	0	0.0	3	10.3	8	27.6	12	41.4	3	10.3	3	10.3
Small Districts Combined[a]	2	7.1	3	10.7	4	14.3	10	35.7	9	32.1	0	0.0
All Districts	12	4.5	16	6.0	46	17.4	111	41.9	70	26.4	10	3.8

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C27. PACE Teachers Survey: Q15c. Please rate your level of agreement with the following statement: I have been able to apply what I've learned from the NH PACE Common Performance Tasks to developing higher quality local performance tasks (i.e., "non-common tasks").

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	6	9.0	3	4.5	11	16.4	32	47.8	13	19.4	2	3.0
Epping	1	4.3	0	0.0	10	43.5	3	13.0	9	39.1	0	0.0
Rochester	6	10.3	4	6.9	10	17.2	25	43.1	9	15.5	4	6.9
Sanborn	0	0.0	2	5.1	5	12.8	10	25.6	21	53.8	1	2.6
Souhegan	0	0.0	0	0.0	2	9.5	5	23.8	14	66.7	0	0.0
White Mountains	2	6.9	5	17.2	4	13.8	13	44.8	1	3.4	4	13.8
Small Districts Combined[a]	1	3.6	4	14.3	5	17.9	7	25.0	9	32.1	2	7.1
All Districts	16	6.0	18	6.8	47	17.7	95	35.8	76	28.7	13	4.9

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C28. PACE Teachers Survey: Q16. Number of local performance tasks (i.e., "non-common tasks") that I plan to administer this year for SCIENCE:

	Zero		1-5		6-10		More than 10	
District	n	%	n	%	n	%	n	%
Concord	35	52.2	27	40.3	4	6.0	1	1.5
Epping	14	60.9	8	34.8	1	4.3	0	0.0
Rochester	34	58.6	20	34.5	3	5.2	1	1.7
Sanborn	23	59.0	13	33.3	3	7.7	0	0.0
Souhegan	12	57.1	4	19.0	4	19.0	1	4.8
White Mountains	15	51.7	12	41.4	1	3.4	1	3.4
Small Districts Combined[a]	10	35.7	18	64.3	0	0.0	0	0.0
All Districts	143	54.0	102	38.5	16	6.0	4	1.5

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C29. PACE Teachers Survey: Q17. Number of local performance tasks (i.e., "non-common tasks") that I plan to administer this year for MATH:

	Zero		1-5		6-10		More than 10	
District	n	%	n	%	n	%	n	%
Concord	30	44.8	31	46.3	6	9.0	0	0.0
Epping	8	34.8	11	47.8	3	13.0	1	4.3
Rochester	27	46.6	26	44.8	4	6.9	1	1.7
Sanborn	16	41.0	23	59.0	0	0.0	0	0.0
Souhegan	13	61.9	6	28.6	2	9.5	0	0.0
White Mountains	12	41.4	13	44.8	4	13.8	0	0.0
Small Districts Combined[a]	10	35.7	14	50.0	2	7.1	2	7.1
All Districts	116	43.8	124	46.8	21	7.9	4	1.5

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C30. PACE Teachers Survey: Q18. Number of local performance tasks (i.e., "non-common tasks") that I plan to administer this year for ELA:

	Zero		1-5		6-10		More than 10	
District	n	%	n	%	n	%	n	%
Concord	33	49.3	30	44.8	4	6.0	0	0.0
Epping	9	39.1	13	56.5	1	4.3	0	0.0
Rochester	24	41.4	23	39.7	7	12.1	4	6.9
Sanborn	17	43.6	16	41.0	6	15.4	0	0.0
Souhegan	14	66.7	6	28.6	1	4.8	0	0.0
White Mountains	11	37.9	10	34.5	7	24.1	1	3.4
Small Districts Combined[a]	10	35.7	13	46.4	4	14.3	1	3.6
All Districts	118	44.5	111	41.9	30	11.3	6	2.3

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C31. PACE Teachers Survey: Q19. Number of local performance tasks (i.e., "non-common tasks") that I plan to administer this year for OTHER content areas (e.g., Social Studies):

	Zero		1-5		6-10		More than 10	
District	n	%	n	%	n	%	n	%
Concord	51	76.1	16	23.9	0	0.0	0	0.0
Epping	16	69.6	7	30.4	0	0.0	0	0.0
Rochester	43	74.1	14	24.1	0	0.0	1	1.7
Sanborn	25	64.1	13	33.3	1	2.6	0	0.0
Souhegan	17	81.0	3	14.3	0	0.0	1	4.8
White Mountains	21	72.4	7	24.1	1	3.4	0	0.0
Small Districts Combined[a]	17	60.7	10	35.7	1	3.6	0	0.0
All Districts	190	71.7	70	26.4	3	1.1	2	0.8

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C32. PACE Teachers Survey: Q20a. Please rate your level of agreement with the following statement: I have received adequate training and preparation to effectively administer the NH PACE Common Performance Tasks.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
District	n	%	n	%	n	%	n	%	n	%
Concord	1	1.5	9	13.4	5	7.5	35	52.2	17	25.4
Epping	1	4.3	3	13.0	5	21.7	6	26.1	8	34.8
Rochester	4	6.9	4	6.9	13	22.4	22	37.9	15	25.9
Sanborn	0	0.0	2	5.1	5	12.8	18	46.2	14	35.90
Souhegan	0	0.0	1	4.8	1	4.8	6	28.6	13	61.9
White Mountains	2	6.9	13	44.8	6	20.7	8	27.6	0	0.0
Small Districts Combined[a]	2	7.1	1	3.6	7	25.0	11	39.3	7	25.0
All Districts	10	3.8	33	12.5	42	15.8	106	40.0	74	27.9

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C33. PACE Teachers Survey: Q20b. Please rate your level of agreement with the following statement: I understand, based on the scaffolding rules, the amount of scaffolding I can employ with the NH PACE Common Performance Tasks.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	14	20.9	7	10.4	34	50.7	11	16.4	1	1.5
Epping	1	4.3	5	21.7	4	17.4	8	34.8	5	21.7	0	0.0
Rochester	2	3.4	10	17.2	11	19.0	23	39.7	11	19.0	1	1.7
Sanborn	1	2.6	5	12.8	3	7.7	18	46.2	12	30.8	0	0.0
Souhegan	0	0.0	2	9.5	0	0.0	8	38.1	11	52.4	0	0.0
White Mountains	4	13.8	5	17.2	6	20.7	10	34.5	1	3.4	3	10.3
Small Districts Combined[a]	2	7.1	4	14.3	5	17.9	13	46.4	4	14.3	0	0.0
All Districts	10	3.8	45	17.0	36	13.6	114	43.0	55	20.8	5	1.9

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C34. PACE Teachers Survey: Q20c. Please rate your level of agreement with the following statement: The NH PACE Implementation Guidelines Manual has been a useful resource in helping me understand how to effectively administer the NH PACE Common Performance Tasks.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	1	1.5	8	11.9	12	17.9	33	49.3	3	4.5	10	14.9
Epping	2	8.7	1	4.3	9	39.1	7	30.4	1	4.3	3	13.0
Rochester	4	6.9	6	10.3	19	32.8	18	31.0	6	10.3	5	8.6
Sanborn	1	2.6	3	7.7	6	15.4	20	51.3	9	23.1	0	0.0
Souhegan	0	0.0	0	0.0	4	19.0	9	42.9	7	33.3	1	4.8
White Mountains	2	6.9	4	13.8	13	44.8	8	27.6	1	3.4	1	3.4
Small Districts Combined[a]	1	3.6	2	7.1	9	32.1	11	39.3	4	14.3	1	3.6
All Districts	11	4.2	24	9.1	72	27.2	106	40.0	31	11.7	21	7.9

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C35. PACE Teachers Survey: Q20d. Please rate your level of agreement with the following statement: The NH PACE Accommodations Guidelines have been useful resources in helping me understand the appropriate accommodations for the NH PACE Common Performance Tasks.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	1	1.5	6	9.0	21	31.3	34	50.7	3	4.5	2	3.0
Epping	1	4.3	3	13.0	8	34.8	10	43.5	0	0.0	1	4.3
Rochester	6	10.3	3	5.2	21	36.2	21	36.2	4	6.9	3	5.2
Sanborn	1	2.6	2	5.1	5	12.8	19	48.7	12	30.8	0	0.0
Souhegan	0	0.0	0	0.0	4	19.0	10	47.6	6	28.6	1	4.8
White Mountains	2	6.9	4	13.8	11	37.9	8	27.6	1	3.4	3	10.3
Small Districts Combined[a]	1	3.6	1	3.6	10	35.7	12	42.9	4	14.3	0	0.0
All Districts	12	4.5	19	7.2	80	30.2	114	43.0	30	11.3	10	3.8

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C36. PACE Teachers Survey: Q20e. Please rate your level of agreement with the following statement: From my perspective, different teachers in different classrooms are administering the same NH PACE Common Performance Tasks in a consistent manner.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	7	10.4	7	10.4	35	52.2	16	23.9	2	3.0	0	0.0
Epping	0	0.0	2	8.7	3	13.0	12	52.2	4	17.4	2	8.7	0	0.0
Rochester	3	5.2	10	17.2	13	22.4	25	43.1	3	5.2	3	5.2	1	1.7
Sanborn	1	2.6	5	12.8	4	10.3	15	38.5	14	35.9	0	0.0	0	0.0
Souhegan	0	0.0	0	0.0	1	4.8	9	42.9	11	52.4	0	0.0	0	0.0
White Mountains	1	3.4	3	10.3	11	37.9	2	6.9	1	3.4	11	37.9	0	0.0
Small Districts Combined[a]	2	7.1	3	10.7	4	14.3	11	39.3	2	7.1	4	14.3	2	7.1
All Districts	7	2.6	30	11.3	43	16.2	109	41.1	51	19.2	22	8.3	3	1.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C37. PACE Teachers Survey: Q21. Based on the characterization of competency-based education, indicate the extent to which competency-based education is integrated into your instruction.

	To No Extent		To a Slight Extent		To Some Extent		To a Great Extent		To a Very Great Extent		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	7	10.4	27	40.3	24	35.8	7	10.4	2	3.0
Epping	0	0.0	2	8.7	11	47.8	8	34.8	2	8.7	0	0.0
Rochester	0	0.0	0	0.0	23	39.7	21	36.2	14	24.1	0	0.0
Sanborn	0	0.0	2	5.1	7	17.9	17	43.6	12	30.8	1	2.6
Souhegan	0	0.0	0	0.0	2	9.5	15	71.4	4	19.0	0	0.0
White Mountains	0	0.0	0	0.0	7	24.1	10	34.5	12	41.4	0	0.0
Small Districts Combined[a]	0	0.0	0	0.0	10	35.7	9	32.1	9	32.1	0	0.0
All Districts	0	0.0	11	4.2	87	32.8	104	39.2	60	22.6	3	1.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C38. PACE Teachers Survey: Q22. I use performance tasks for instructional purposes:

	Never		A Couple Times a Month or Less		Approximately Once Each Week		2-3 Times Each Week		Nearly Every Day	
District	n	%	n	%	n	%	n	%	n	%
Concord	8	11.9	36	53.7	18	26.9	5	7.5	0	0.0
Epping	1	4.3	16	69.6	5	21.7	1	4.3	0	0.0
Rochester	3	5.2	28	48.3	17	29.3	8	13.8	2	3.4
Sanborn	4	10.3	20	51.3	6	15.4	8	20.5	1	2.6
Souhegan	0	0.0	8	38.1	8	38.1	3	14.3	2	9.5
White Mountains	2	6.9	14	48.3	9	31.0	3	10.3	1	3.4
Small Districts Combined[a]	2	7.4	15	55.6	6	22.2	3	11.1	1	3.7
All Districts	20	7.6	137	51.9	69	26.1	31	11.7	7	2.7

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C39. PACE Teachers Survey: Q23. I use performance tasks for assessment purposes:

	Never		A Couple Times a Month or Less		Approximately Once Each Week		2-3 Times Each Week		Nearly Every Day	
District	n	%	n	%	n	%	n	%	n	%
Concord	4	6.0	52	77.6	9	13.4	1	1.5	1	1.5
Epping	1	4.3	21	91.3	0	0.0	1	4.3	0	0.0
Rochester	6	10.3	37	63.8	14	24.1	1	1.7	0	0.0
Sanborn	2	5.1	30	76.9	5	12.8	1	2.6	1	2.6
Souhegan	0	0.0	14	66.7	4	19.0	2	9.5	1	4.8
White Mountains	2	6.9	20	69.0	2	6.9	4	13.8	1	3.4
Small Districts Combined[a]	2	7.4	23	85.2	2	7.4	0	0.0	0	0.0
All Districts	17	6.4	197	74.6	36	13.6	10	3.8	4	1.5

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C40. PACE Teachers Survey: Q24a1. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on instructional practice, such that instruction occurs at a higher DOK in my classroom.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	4	6.0	8	11.9	36	53.7	17	25.4	2	3.0
Epping	0	0.0	1	4.3	2	8.7	12	52.2	8	34.8	0	0.0
Rochester	2	3.4	1	1.7	10	17.2	30	51.7	14	24.1	1	1.7
Sanborn	0	0.0	3	7.7	4	10.3	14	35.9	17	43.6	1	2.6
Souhegan	0	0.0	0	0.0	2	9.5	9	42.9	10	47.6	0	0.0
White Mountains	0	0.0	2	6.9	2	6.9	16	55.2	9	31.0	0	0.0
Small Districts Combined[a]	0	0.0	0	0.0	3	11.1	15	55.6	8	29.6	1	3.7
All Districts	2	0.8	11	4.2	31	11.7	132	50.0	83	31.4	5	1.9

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C41. PACE Teachers Survey: Q24a2. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on instructional practice, such that instruction occurs at a higher DOK in my school.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	4	6.0	8	11.9	40	59.7	7	10.4	8	11.9
Epping	0	0.0	1	4.3	2	8.7	14	60.9	2	8.7	4	17.4
Rochester	2	3.4	1	1.7	12	20.7	30	51.7	5	8.6	8	13.8
Sanborn	0	0.0	3	7.7	8	20.5	16	41.0	9	23.1	3	7.7
Souhegan	0	0.0	0	0.0	2	9.5	8	38.1	9	42.9	2	9.5
White Mountains	0	0.0	1	3.4	7	24.1	10	34.5	6	20.7	5	17.2
Small Districts Combined[a]	0	0.0	1	3.7	2	7.4	14	51.9	4	14.8	6	22.2
All Districts	2	0.8	11	4.2	41	15.5	132	50.0	42	15.9	36	13.6

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C42. PACE Teachers Survey: Q24b1. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on student engagement while completing performance tasks in my classroom.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	3	4.5	1	1.5	7	10.4	35	52.2	21	31.3	0	0.0
Epping	0	0.0	1	4.3	7	30.4	4	17.4	10	43.5	1	4.3
Rochester	2	3.4	2	3.4	11	19.0	29	50.0	13	22.4	1	1.7
Sanborn	1	2.6	2	5.1	5	12.8	12	30.8	19	48.7	0	0.0
Souhegan	0	0.0	0	0.0	0	0.0	9	42.9	12	57.1	0	0.0
White Mountains	0	0.0	2	6.9	5	17.2	13	44.8	9	31.0	0	0.0
Small Districts Combined[a]	1	3.7	0	0.0	2	7.4	18	66.7	5	18.5	1	3.7
All Districts	7	2.7	8	3.0	37	14.0	120	45.5	89	33.7	3	1.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C43. PACE Teachers Survey: Q24b2. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on student engagement while completing performance tasks in my school.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	1	1.5	2	3.0	8	11.9	37	55.2	9	13.4	10	14.9
Epping	0	0.0	1	4.3	6	26.1	6	26.1	6	26.1	4	17.4
Rochester	2	3.4	2	3.4	10	17.2	29	50.0	8	13.8	7	12.1
Sanborn	0	0.0	3	7.7	7	17.9	14	35.9	13	33.3	2	5.1
Souhegan	0	0.0	0	0.0	2	9.5	6	28.6	11	52.4	2	9.5
White Mountains	0	0.0	3	10.3	7	24.1	8	27.6	6	20.7	5	17.2
Small Districts Combined[a]	0	0.0	1	3.7	0	0.0	16	59.3	3	11.1	7	25.9
All Districts	3	1.1	12	4.5	40	15.2	116	43.9	56	21.2	37	14.0

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C44. PACE Teachers Survey: Q24c1. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on student engagement in learning overall in my classroom.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	3	4.5	15	22.4	30	44.8	18	26.9	1	1.5
Epping	0	0.0	2	8.7	6	26.1	8	34.8	7	30.4	0	0.0
Rochester	3	5.2	2	3.4	14	24.1	29	50.0	9	15.5	1	1.7
Sanborn	0	0.0	1	2.6	9	23.1	12	30.8	17	43.6	0	0.0
Souhegan	0	0.0	0	0.0	2	9.5	10	47.6	9	42.9	0	0.0
White Mountains	0	0.0	2	6.9	5	17.2	14	48.3	8	27.6	0	0.0
Small Districts Combined[a]	0	0.0	1	3.7	3	11.1	16	59.3	6	22.2	1	3.7
All Districts	3	1.1	11	4.2	54	20.5	119	45.1	74	28.0	3	1.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C45. PACE Teachers Survey: Q24c2. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on student engagement in learning overall in my school.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	3	4.5	11	16.4	35	52.2	9	13.4	9	13.4
Epping	0	0.0	1	4.3	7	30.4	9	39.1	3	13.0	3	13.0
Rochester	4	6.9	1	1.7	16	27.6	26	44.8	4	6.9	7	12.1
Sanborn	0	0.0	2	5.1	10	25.6	14	35.9	11	28.2	2	5.1
Souhegan	0	0.0	0	0.0	3	14.3	7	33.3	9	42.9	2	9.5
White Mountains	0	0.0	2	6.9	6	20.7	10	34.5	6	20.7	5	17.2
Small Districts Combined[a]	0	0.0	1	3.7	3	11.1	13	48.1	4	14.8	6	22.2
All Districts	4	1.5	10	3.8	56	21.2	114	43.2	46	17.4	34	12.9

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C46. PACE Teachers Survey: Q25a. Please rate your level of agreement with the following statement: The scoring rubrics for the NH PACE Common Performance Tasks are sufficiently clear and detailed to ensure that separate scorers scoring the same student work arrive at the same score.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	6	9.0	11	16.4	14	20.9	30	44.8	5	7.5	1	1.5
Epping	3	13.0	5	21.7	5	21.7	10	43.5	0	0.0	0	0.0
Rochester	9	15.5	8	13.8	24	41.4	12	20.7	2	3.4	3	5.2
Sanborn	2	5.1	3	7.7	7	17.9	20	51.3	7	17.9	0	0.0
Souhegan	0	0.0	1	4.8	5	23.8	14	66.7	1	4.8	0	0.0
White Mountains	0	0.0	5	17.2	7	24.1	10	34.5	2	6.9	5	17.2
Small Districts Combined[a]	2	7.4	8	29.6	1	3.7	14	51.9	2	7.4	0	0.0
All Districts	22	8.3	41	15.5	63	23.9	110	41.7	19	7.2	9	3.4

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C47. PACE Teachers Survey: Q25b. Please rate your level of agreement with the following statement: The scoring resources available on the LibGuide effectively explain how to score the student work on the NH PACE Common Performance Tasks.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	1	1.5	4	6.0	18	26.9	22	32.8	3	4.5	19	28.4
Epping	1	4.3	3	13.0	11	47.8	5	21.7	0	0.0	3	13.0
Rochester	5	8.6	9	15.5	21	36.2	13	22.4	0	0.0	10	17.2
Sanborn	1	2.6	1	2.6	9	23.1	18	46.2	7	17.9	3	7.7
Souhegan	0	0.0	0	0.0	3	14.3	10	47.6	2	9.5	6	28.6
White Mountains	0	0.0	3	10.3	9	31.0	9	31.0	2	6.9	6	20.7
Small Districts Combined[a]	2	7.4	1	3.7	7	25.9	14	51.9	1	3.7	2	7.4
All Districts	10	3.8	21	8.0	78	29.5	91	34.5	15	5.7	49	18.6

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C48. PACE Teachers Survey: Q25c. Please rate your level of agreement with the following statement: The within-district double scoring is effective in ensuring the reliability of scores on the NH PACE Common Performance Tasks.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	1	1.5	2	3.0	13	19.4	38	56.7	8	11.9	5	7.5
Epping	1	4.3	2	8.7	4	17.4	8	34.8	7	30.4	1	4.3
Rochester	4	6.9	2	3.4	11	19.0	30	51.7	8	13.8	3	5.2
Sanborn	0	0.0	1	2.6	3	7.7	19	48.7	16	41.0	0	0.0
Souhegan	0	0.0	1	4.8	1	4.8	9	42.9	10	47.6	0	0.0
White Mountains	0	0.0	1	3.4	10	34.5	4	13.8	1	3.4	13	44.8
Small Districts Combined[a]	1	3.7	1	3.7	2	7.4	19	70.4	4	14.8	0	0.0
All Districts	7	2.7	10	3.8	44	16.7	127	48.1	54	20.5	22	8.3

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table C49. PACE Teachers Survey: Q26. Please rate your level of agreement with the following statement: The NH PACE Common Performance Tasks are more accessible to a greater range of student learning needs (e.g., students with disabilities, English language learners) than traditional standardized tests.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	8	11.9	11	16.4	14	20.9	25	37.3	8	11.9	1	1.5
Epping	2	8.7	4	17.4	8	34.8	7	30.4	2	8.7	0	0.0
Rochester	5	8.6	10	17.2	18	31.0	18	31.0	3	5.2	4	6.9
Sanborn	1	2.6	5	12.8	9	23.1	11	28.2	13	33.3	0	0.0
Souhegan	2	9.5	1	4.8	2	9.5	8	38.1	8	38.1	0	0.0
White Mountains	4	13.8	4	13.8	8	27.6	6	20.7	2	6.9	5	17.2
Small Districts Combined[a]	4	14.8	3	11.1	10	37.0	5	18.5	5	18.5	0	0.0
All Districts	26	9.8	38	14.4	69	26.1	80	30.3	41	15.5	10	3.8

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Appendix D: Other Teacher Survey Response Frequencies

Table D1. Other Teachers Survey: Q2a. Please indicate the grade level(s) you currently teach. Select all that apply.

	Kindergarten		Grade 1		Grade 2		Grade 3		Grade 4		Grade 5	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	6	14.6	8	19.5	8	19.5	8	19.5	7	17.1	7	17.1
Epping	6	18.2	8	24.2	8	24.2	4	12.1	3	9.1	4	12.1
Rochester	5	10.6	11	23.4	11	23.4	4	8.5	5	10.6	4	8.5
Sanborn	11	20.4	10	18.5	10	18.5	9	16.7	10	18.5	9	16.7
Souhegan	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
White Mountains	15	37.5	17	42.5	16	40.0	12	30.0	12	30.0	15	37.5
Small Districts Combined[a]	11	37.9	10	34.5	12	41.4	6	20.7	4	13.8	4	13.8
All Districts	54	20.8	64	24.6	65	25.0	43	16.5	41	15.8	43	16.5

[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D2. Other Teachers Survey: Q2b. Please indicate the grade level(s) you currently teach. Select all that apply.

	Grade 6		Grade 7		Grade 8		Grade 9		Grade 10		Grade 11		Grade 12	
District	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Concord	6	14.6	10	24.4	10	24.4	8	19.5	10	24.4	14	34.1	12	29.3
Epping	2	6.1	2	6.1	7	21.2	7	21.2	8	24.2	11	33.3	10	30.3
Rochester	7	14.9	7	14.9	6	12.8	13	27.7	13	27.7	15	31.9	13	27.7
Sanborn	8	14.8	9	16.7	9	16.7	10	18.5	13	24.1	17	31.9	17	31.5
Souhegan	0	0.0	0	0.0	0	0.0	10	62.5	11	68.8	14	87.5	14	87.5
White Mountains	14	35.0	11	27.5	13	32.5	11	27.5	11	27.5	12	30.0	12	30.0
Small Districts Combined[a]	4	13.8	5	17.2	5	17.2	9	31.0	7	24.1	7	24.1	7	24.1
All Districts	41	15.8	44	16.9	50	19.2	68	26.2	73	28.1	90	34.6	85	32.7

[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D3. Other Teachers Survey: Q3a. Please indicate the content area(s) for which you currently teach. Select all that apply.

	ELA		Mathematics		Science		Art		Music/ Fine Arts		Social Studies/ History		World Language	
District	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Concord	7	17.1	9	22.0	5	12.2	5	12.2	1	2.4	6	14.6	3	7.3
Epping	12	36.4	15	45.5	10	30.3	2	6.1	1	3.0	13	39.4	3	9.1
Rochester	21	44.7	21	44.7	20	42.6	5	10.6	1	2.1	19	40.4	0	0.0
Sanborn	21	38.9	16	29.6	16	29.6	4	7.4	2	3.7	19	35.2	4	7.4
Souhegan	5	31.3	1	6.3	3	18.8	0	0.0	0	0.0	5	31.3	3	18.8
White Mountains	17	42.5	16	40.0	10	25.0	3	7.5	2	5.0	14	35.0	4	10.0
Small Districts Combined[a]	14	48.3	13	44.8	12	41.4	4	13.8	5	17.2	11	37.9	0	0.0
All Districts	98	37.7	91	35.0	76	29.2	23	8.8	12	4.6	87	33.5	17	6.5

Note. ELA= English-Language Arts

[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D4. Other Teachers Survey: Q3b. Please indicate the content area(s) for which you currently teach. Select all that apply.

	Special Education		Vocational Studies		Physical Education/ Health		Library		Drama		Computers/ Technology		Other	
District	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Concord	2	4.9	4	9.8	4	9.8	2	4.9	0	0.0	4	9.8	9	22.0
Epping	0	0.0	0	0.0	1	3.0	0	0.0	0	0.0	0	0.0	6	18.2
Rochester	3	6.4	3	6.4	3	6.4	1	2.1	0	0.0	6	12.8	11	23.4
Sanborn	2	3.7	1	1.9	2	3.7	1	1.9	0	0.0	2	3.7	5	9.3
Souhegan	2	12.5	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	3	18.8
White Mountains	0	0.0	2	5.0	2	5.0	1	2.5	0	0.0	3	7.5	4	10.0
Small Districts Combined[a]	2	6.9	0	0.0	1	3.4	3	10.3	1	3.4	4	13.8	5	17.2
All Districts	12	4.6	10	3.8	13	5.0	8	3.1	1	0.4	19	7.3	23	8.8

[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D5. Other Teachers Survey: Q4. Have you switched schools or districts since the 2014–15 school year? Select all that apply.

	No		Yes, I switched schools and/or districts in 2014-15		Yes, I switched schools and/or districts in 2015-16		Yes, I switched schools and/or districts in 2016-17	
District	n	%	n	%	n	%	n	%
Concord	35	85.4	1	2.4	2	4.9	4	9.8
Epping	26	78.8	0	0.0	2	6.1	5	15.2
Rochester	40	85.1	3	6.4	2	4.3	2	4.3
Sanborn	44	81.5	2	3.7	4	7.4	7	13.0
Souhegan	14	87.5	1	6.3	0	0.0	1	6.3
White Mountains	38	95.0	1	2.5	0	0.0	1	2.5
Small Districts Combined[a]	22	75.9	1	3.4	3	10.3	3	10.3
All Districts	219	84.2	9	3.5	13	5.0	23	8.8

[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D6. Other Teachers Survey: Q5. Please rate your level of familiarity with the NH Performance Assessment of Competency Education (PACE) pilot program at your school.

	Unfamiliar		Somewhat Unfamiliar		Somewhat Familiar		Very Familiar		Extremely Familiar	
District	n	%	n	%	n	%	n	%	n	%
Concord	6	14.6	8	19.5	19	46.3	7	17.1	1	2.4
Epping	4	12.1	6	18.2	18	54.5	5	15.2	5	0.0
Rochester	7	14.9	7	14.9	21	44.7	10	21.3	2	4.3
Sanborn	2	3.7	1	1.9	33	61.1	14	25.9	4	7.4
Souhegan	0	0.0	0	0.0	5	31.3	8	50.0	3	18.8
White Mountains	0	0.0	7	17.5	27	67.5	6	15.0	0	0.0
Small Districts Combined[a]	3	10.3	2	6.9	15	51.7	8	27.6	1	3.4
All Districts	22	8.5	31	11.9	138	53.1	58	22.3	11	4.2

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D7. Other Teachers Survey: Q6. Do you develop and/or administer local performance tasks (i.e., "non-common tasks") in the content area(s)/grade(s) you teach?

	Yes		No	
District	n	%	n	%
Concord	22	53.7	19	46.3
Epping	22	66.7	11	33.3
Rochester	29	61.7	18	38.3
Sanborn	45	83.3	9	16.7
Souhegan	12	75.0	4	25.0
White Mountains	24	60.0	16	40.0
Small Districts Combined[a]	19	65.5	10	34.5
All Districts	173	66.5	87	33.5

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D8. Other Teachers Survey: Q7. Please indicate the approximate number of local performance tasks (i.e., "non-common tasks") that you plan to administer this academic year.

	Zero		1–5		6–10		More than 10	
District	n	%	n	%	n	%	n	%
Concord	16	39.0	22	53.7	2	4.9	1	2.4
Epping	7	21.2	15	45.5	5	15.2	6	18.2
Rochester	15	31.9	23	48.9	4	8.5	5	10.6
Sanborn	9	16.7	28	51.9	13	24.1	4	7.4
Souhegan	4	25.0	4	25.0	5	31.3	3	18.8
White Mountains	7	17.5	21	52.5	6	15.0	6	15.0
Small Districts Combined[a]	9	31.0	17	58.6	1	3.4	2	6.9
All Districts	67	25.8	130	50.0	36	13.8	27	10.4

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D9. Other Teachers Survey: Q8. Please rate your level of agreement with the following statement: My school's administration (e.g., principal, assistant principal, curriculum director) is supportive of the PACE initiative.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	0	0.0	4	9.8	15	36.6	16	39.0	6	14.6
Epping	1	3.0	0	0.0	2	6.1	17	51.5	10	30.3	3	9.1
Rochester	1	2.1	0	0.0	3	6.4	14	29.8	20	42.6	9	19.1
Sanborn	2	3.7	0	0.0	1	1.9	16	29.6	33	61.1	2	3.7
Souhegan	1	6.3	0	0.0	0	0.0	1	6.3	14	87.5	0	0.0
White Mountains	0	0.0	0	0.0	2	5.0	14	35.0	23	57.5	1	2.5
Small Districts Combined[a]	0	0.0	1	3.4	2	6.9	10	34.5	16	55.2	0	0.0
All Districts	5	1.9	1	0.4	14	5.4	87	33.5	132	50.8	21	8.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D10. Other Teachers Survey: Q9. Think about the teachers at your school who are familiar with the PACE initiative; this includes teachers who administer NH PACE common performance tasks AND teachers who do not administer common performance tasks, but who are familiar with the PACE initiative via attendance at faculty meetings, informal discussions with teachers administering PACE tasks, etc. Please select the statement that most closely reflects your perception of their opinion about PACE.

	They all have a favorable opinion of the PACE initiative.		Most have a favorable opinion of the PACE initiative		Some have a favorable opinion of the PACE initiative		Few have a favorable opinion of the PACE initiative		None have a favorable opinion of the PACE initiative		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	2	4.9	12	29.3	13	31.7	2	4.9	0	0.0	12	29.3
Epping	1	3.0	6	18.2	12	36.4	10	30.3	0	0.0	4	12.1
Rochester	2	4.3	11	23.4	16	34.0	7	14.9	1	2.1	10	21.3
Sanborn	5	9.3	22	40.7	17	31.5	3	5.6	0	0.0	7	13.0
Souhegan	1	6.3	13	81.3	0	0.0	0	0.0	1	6.3	1	6.3
White Mountains	2	5.0	11	27.5	14	35.0	6	15.0	0	0.0	7	17.5
Small Districts Combined[a]	2	6.9	10	34.5	8	27.6	4	13.8	0	0.0	5	17.2
All Districts	15	5.8	85	32.7	80	30.8	32	12.3	2	0.8	46	17.7

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D11. Other Teachers Survey: Q10. Based on the characterization of competency-based education, indicate the extent to which competency-based education is integrated into your instruction.

	To No Extent		To a Slight Extent		To Some Extent		To a Great Extent		To a Very Great Extent		Not Applicable	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	2	4.9	3	7.3	13	31.7	14	34.1	7	17.1	2	4.9
Epping	0	0.0	1	3.0	17	51.5	13	39.4	2	6.1	0	0.0
Rochester	0	0.0	3	6.4	17	36.2	13	27.7	13	27.7	1	2.1
Sanborn	0	0.0	1	1.9	9	16.7	24	44.4	19	35.2	1	1.9
Souhegan	0	0.0	0	0.0	1	6.3	7	43.8	6	37.5	2	12.5
White Mountains	0	0.0	0	0.0	8	20.0	19	47.5	12	30.0	1	2.5
Small Districts Combined[a]	0	0.0	1	3.4	9	31.0	7	24.1	12	41.4	0	0.0
All Districts	2	0.8	9	3.5	74	28.5	97	37.3	71	27.3	7	27.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D12. Other Teachers Survey: Q11. I use performance tasks for instructional purposes:

	Never		A Couple Times a Month or Less		Approximately Once Each Week		2-3 Times Each Week		Nearly Every Day	
District	n	%	n	%	n	%	n	%	n	%
Concord	6	14.6	17	41.5	8	19.5	6	14.6	4	9.8
Epping	3	9.1	17	51.5	7	21.2	2	6.1	4	12.1
Rochester	3	6.4	22	46.8	7	14.9	7	14.9	8	17.0
Sanborn	4	7.5	22	41.5	13	24.5	6	11.3	8	15.1
Souhegan	3	18.8	6	37.5	3	18.8	2	12.5	2	12.5
White Mountains	2	5.0	13	32.5	4	10.0	12	30.0	9	22.5
Small Districts Combined[a]	2	6.9	11	37.9	8	27.6	3	10.3	5	17.2
All Districts	23	8.9	108	41.7	50	19.3	38	14.7	40	15.4

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D13. Other Teachers Survey: Q12. I use performance tasks for assessment purposes:

	Never		A Couple Times a Month or Less		Approximately Once Each Week		2-3 Times Each Week		Nearly Every Day	
District	n	%	n	%	n	%	n	%	n	%
Concord	7	17.1	21	51.2	5	12.2	5	12.2	3	7.3
Epping	4	12.1	22	66.7	5	15.2	2	6.1	0	0.0
Rochester	5	10.6	25	53.2	7	14.9	4	8.5	6	12.8
Sanborn	2	3.8	35	66.0	8	15.1	2	3.8	6	11.3
Souhegan	2	12.5	10	62.5	2	12.5	1	6.3	1	6.3
White Mountains	2	5.0	21	52.5	7	17.5	7	17.5	3	7.5
Small Districts Combined[a]	2	6.9	14	48.3	7	24.1	4	13.8	2	6.9
All Districts	24	9.3	148	57.1	41	15.8	25	9.7	21	8.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D14. Other Teachers Survey: Q13a1. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on instructional practice, such that instruction occurs at a higher DOK in my classroom.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	1	2.4	7	17.1	14	34.1	16	39.0	3	7.3
Epping	0	0.0	1	3.0	6	18.2	18	54.5	8	24.2	0	0.0
Rochester	0	0.0	0	0.0	8	17.0	25	53.2	12	25.5	2	4.3
Sanborn	2	3.8	0	0.0	8	15.4	22	42.3	19	36.5	1	1.9
Souhegan	0	0.0	0	0.0	0	0.0	7	46.7	6	40.0	2	13.3
White Mountains	0	0.0	0	0.0	1	2.5	22	55.0	16	40.0	1	2.5
Small Districts Combined[a]	0	0.0	0	0.0	1	3.4	11	37.9	17	58.6	0	0.0
All Districts	2	0.8	2	0.8	31	12.1	119	46.3	94	36.6	9	3.5

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D15. Other Teachers Survey: Q13a2. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on instructional practice, such that instruction occurs at a higher DOK in my school.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	0	0.0	7	17.1	18	43.9	6	14.6	10	24.4
Epping	0	0.0	1	3.0	9	27.3	16	48.5	6	18.2	1	3.0
Rochester	0	0.0	2	4.3	8	17.0	23	48.9	7	14.9	7	14.9
Sanborn	2	3.8	0	0.0	6	11.5	20	38.5	19	36.5	5	9.6
Souhegan	1	6.7	0	0.0	0	0.0	7	46.7	6	40.0	1	6.7
White Mountains	0	0.0	0	0.0	3	7.5	15	37.5	11	27.5	11	27.5
Small Districts Combined[a]	0	0.0	0	0.0	0	0.0	16	55.2	12	41.4	1	3.4
All Districts	3	1.2	3	1.2	33	12.8	115	44.7	67	26.1	36	14.0

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D16. Other Teachers Survey: Q13b1. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on student engagement while completing performance tasks in my classroom.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	0	0.0	6	14.6	16	39.0	17	41.5	2	4.9
Epping	1	3.0	1	3.0	9	27.3	16	48.5	5	15.2	1	3.0
Rochester	0	0.0	0	0.0	7	14.9	17	36.2	20	42.6	3	6.4
Sanborn	3	5.8	0	0.0	7	13.5	19	36.5	22	42.3	1	1.9
Souhegan	1	6.7	0	0.0	0	0.0	7	46.7	7	46.7	0	0.0
White Mountains	0	0.0	0	0.0	3	7.5	22	55.0	14	35.0	1	2.5
Small Districts Combined[a]	0	0.0	0	0.0	3	10.3	16	55.2	10	34.5	0	0.0
All Districts	5	1.9	1	0.4	35	13.6	113	44.0	95	37.0	8	3.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D17. Other Teachers Survey: Q13b2. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on student engagement while completing performance tasks in my school.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	0	0.0	6	14.6	17	41.5	7	17.1	11	26.8
Epping	0	0.0	2	6.1	10	30.3	15	45.5	3	9.1	3	9.1
Rochester	0	0.0	2	4.3	8	17.0	18	38.3	11	23.4	8	17.0
Sanborn	2	3.8	0	0.0	5	9.6	21	40.4	19	36.5	5	9.6
Souhegan	0	0.0	0	0.0	0	0.0	6	40.0	6	40.0	3	20.0
White Mountains	0	0.0	0	0.0	3	7.5	16	40.0	10	25.0	11	27.5
Small Districts Combined[a]	0	0.0	0	0.0	2	6.9	18	62.1	6	20.7	3	10.3
All Districts	2	0.8	4	1.6	34	13.2	111	43.2	62	24.1	44	17.1

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D18. Other Teachers Survey: Q13c1. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on student engagement in learning overall in my classroom.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	0	0.0	8	19.5	15	36.6	16	39.0	2	4.9
Epping	0	0.0	1	3.0	13	39.4	10	30.3	8	24.2	1	3.0
Rochester	0	0.0	1	2.1	7	14.9	20	42.6	16	34.0	3	6.4
Sanborn	3	5.8	0	0.0	8	15.4	23	44.2	17	32.7	1	1.9
Souhegan	0	0.0	0	0.0	0	0.0	7	46.7	7	46.7	1	6.7
White Mountains	0	0.0	0	0.0	2	5.0	25	62.5	12	30.0	1	2.5
Small Districts Combined[a]	0	0.0	1	3.4	2	6.9	15	51.7	11	37.9	0	0.0
All Districts	3	1.2	3	1.2	40	15.6	115	44.7	87	33.9	9	3.5

Note. Numbers and percentages are based on the total number of valid responses.
[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Table D19. Other Teachers Survey: Q13c2. Please rate your level of agreement with the following statement: Implementing performance tasks has had a positive impact on student engagement in learning overall in my school.

	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		Don't Know	
District	n	%	n	%	n	%	n	%	n	%	n	%
Concord	0	0.0	0	0.0	7	17.1	16	39.0	6	14.6	12	29.3
Epping	0	0.0	0	0.0	15	45.5	10	30.3	6	18.2	2	6.1
Rochester	0	0.0	2	4.3	10	21.3	18	38.3	9	19.1	8	17.0
Sanborn	2	3.8	0	0.0	6	11.5	22	42.3	18	34.6	4	7.7
Souhegan	1	6.7	0	0.0	0	0.0	6	40.0	6	40.0	2	13.3
White Mountains	0	0.0	0	0.0	3	7.5	17	42.5	9	22.5	11	27.5
Small Districts Combined[a]	0	0.0	1	3.4	1	3.4	19	65.5	6	20.7	2	6.9
All Districts	3	1.2	3	1.2	42	16.3	108	42.0	60	23.3	41	16.0

Note. Numbers and percentages are based on the total number of valid responses.

[a] Represents the Monroe, Seacoast, and Pittsfield school districts.

Appendix E: 2016 PACE Results with No 2015 Comparison

The body of this report provides an analysis comparing 2015 PACE results to 2016 PACE results. The figures in this appendix depict high school and 2016 results for which we do not have 2015 comparison points.

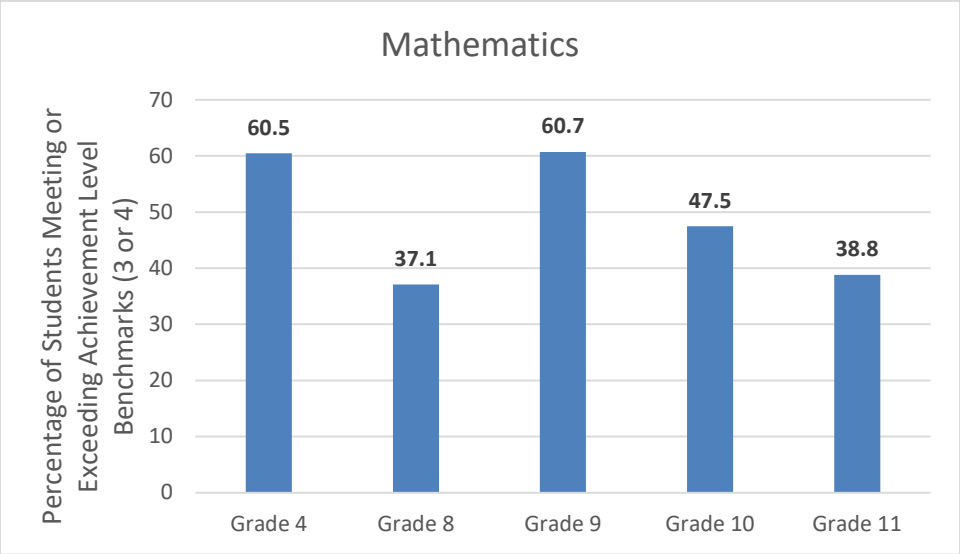


Figure E1. Percentages of students meeting or exceeding achievement level benchmarks in 2016 Mathematics, grades 4, 8, 9, 10, and 11.

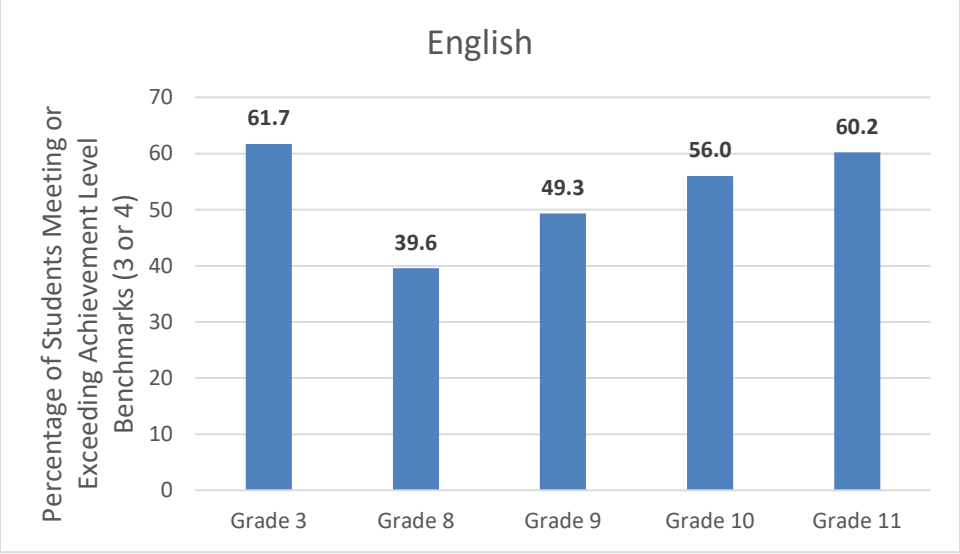


Figure E2. Percentages of students meeting or exceeding achievement level benchmarks in 2016 ELA, grades 3, 8, 9, 10, and 11.

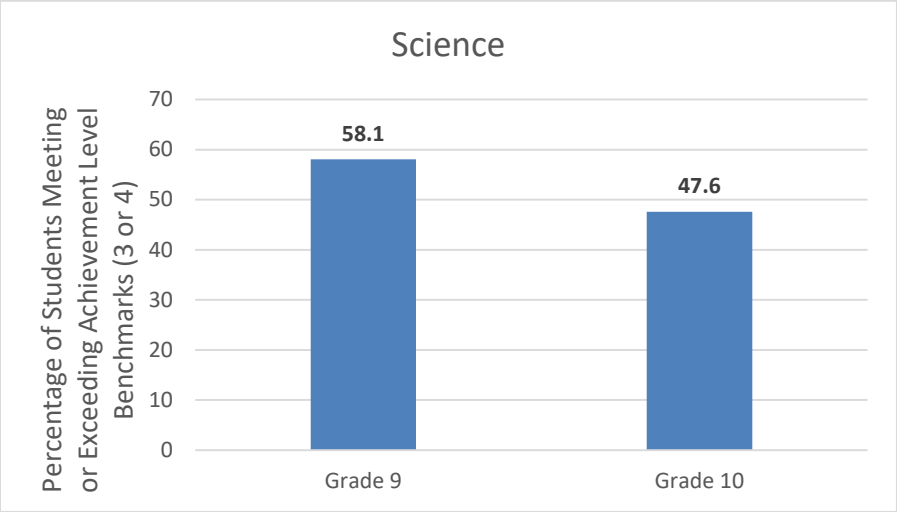


Figure E1. Percentages of students meeting or exceeding achievement level benchmarks in 2016 Science, grades 9 and 10.