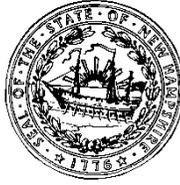


Virginia M. Barry, Ph.D.
Commissioner of Education
Tel. 603-271-3144



Paul K. Leather
Deputy Commissioner of Education
603-271-3801

STATE OF NEW HAMPSHIRE
DEPARTMENT OF EDUCATION
101 Pleasant Street
Concord, N.H. 03301-3860
FAX 603-271-1953
Citizens Services Line 1-800-339-9900
www.education.nh.gov

NEW HAMPSHIRE

PERFORMANCE ASSESSMENT OF COMPETENCY

EDUCATION:

AN ACCOUNTABILITY PILOT PROPOSAL

TO THE

UNITED STATES DEPARTMENT OF EDUCATION

November 21, 2014

Table of Contents

Introduction and Rationale..... 3

History of and Rationale for the PACE Option 4

 Rationale for PACE 5

Implementation Plan 6

 Requirements for participating districts (“guardrails”)..... 6

 The Task Bank 8

 Technical and professional learning support 10

 Benchmarks for success 13

 Communication..... 14

Technical Issues and Considerations 15

 Comparability of annual determinations..... 15

 Equity 17

 Student Progress and Educator Effectiveness..... 17

 Research and Development..... 18

 Utility 18

Appendices..... 19

Introduction and Rationale

The New Hampshire Department of Education (NHDOE) is pleased to provide the following information as a follow-up to questions raised by USED leadership regarding our proposed pilot accountability system, the Performance Assessment of Competency Education or PACE. This memo is our formal request to allow NH DOE to waive a limited amount of annual state-level achievement testing for four NH school districts and to base accountability determinations on a combination of local, common, and state-level assessments.

New Hampshire is committed to ensuring that all students graduate from its schools college and career-ready. Although New Hampshire is one of the highest performing states in the country and has been improving its performance over the last 15 years, the State is not satisfied with the current levels of school and student performance. A key factor contributing to this unease with the status quo stems from the unacceptably high level of remediation required by students entering post-secondary institutions. But another cause of our motivation to improve stems from knowing that we can do more to engage all students in meaningful learning opportunities. NH educational leaders argue that we are beginning to “top out” on the level of performance that can be expected in a top-down or externally-controlled accountability system. Rather than continue to operate within such a system, education leaders in New Hampshire want to shift to more internally-focused improvement system aligned with research on human and organizational learning and improvement. In collaboration with a broad range of stakeholders throughout the state, the New Hampshire Department of Education (NHDOE) is developing an expanded view of assessment and accountability, grounded in research that it proposes to pilot in a small number of districts in the 2014-2015 school year. There are several key components to this approach that NH DOE believes will help them achieve better results for all students:

- ✓ Explicit involvement of local educational leaders in designing and implementing the accountability system,
- ✓ Intense and reciprocal support on behalf of the NHDOE for local districts involved in this initiative that will include technical, policy, and practical guidance,
- ✓ Use of a competency-based approach to instruction, learning, and assessment which can best support the goal of significant improvements in college and career readiness, and
- ✓ Use of authentic, instructionally-relevant, and validated performance-based assessments, alongside periodic administration of Smarter Balanced assessments of state standards in math and ELA, for the purpose of tracking and reporting the progress of students, schools, districts, and educators.

This new approach does not change the state’s firm commitment to accountability for the purposes of improving student learning and outcomes, especially for educationally disadvantaged student groups, as well as supporting high quality educator, leader, and school support and evaluation systems. However, the State argues that an improvement-focused approach improves how the state collects and uses information to better meet the needs of educators and students in New Hampshire. We present details of this request in this memo, organized in three main sections:

1. History of and rationale for the PACE option
2. Implementation plan
3. Technical issues

History of and Rationale for the PACE Option

In spring 2012, Commissioner Barry and Deputy Commissioner Leather met with Secretary Duncan and staff to discuss a new kind of school accountability system based on competency-based education. In July, 2012, a concept paper (NHDOE, 2012) describing a new theory of action for accountability based on an expanded definition of college and career readiness and competency education for students and educators was created, conceptualizing a new model of accountability (see Appendix A). However, upon examination, the NH Department of Education, in conjunction with NH stakeholders identified that the moving parts of such a new system had yet to be fully conceptualized, let alone constructed. For this reason, a first ESEA waiver request was submitted that was based on the current system of accountability, but with the provision that NH would start to build the new system. That same year, the NHDOE began a large scale professional development initiative with teams of NH educators from a first cohort of schools who had dedicated themselves to implementing competency education, K-12.

With the generous support from the Nellie Mae Education Foundation, the NHDOE in 2013 constructed a performance assessment model of local accountability to support the implementation of competency education based on the tenets of a white paper completed by the Accountability 3.0 Taskforce of the Innovation Lab Network of the Council of Chief State School Officers, (CCSSO). This model, which laid the foundation for the PACE proposal (see Appendix B), conceptualized a scalable model of state and local accountability supported by common performance assessments juried at the state level and aligned to NH state level graduation competencies in English language arts, mathematics, science, and work study practices. In the spring of the 2014, with the continued support of the Nellie Mae Education Foundation (NMEF) and the William and Flora Hewlett Foundation, the NHDOE established a pilot of Performance Assessment of Competency Education, (PACE) comprised of four (4) implementing districts and four (4) planning districts dedicated to fully develop and implement a system that would satisfy the accountability expectations of a federal system for the 2014-2015 school year. Significant partners in this work included the Center for Collaborative Education (CCE) and the National Center for the Improvement of Educational Assessment (Center for Assessment), charged with assuring a rigorous valid and reliable system of common system of complex performance assessments aligned with grade span summative Smarter Balanced Assessment Consortium Assessments. Working with support from CCSSO, NHDOE, key stakeholders, and project partners from the Center for Assessment and CCE began the detailed work of framing the specifics of the PACE proposal in early 2014. Since that time, a first draft of a PACE Guide for implementing districts has been completed, (see Appendix C), that outlines the parameters of the learning, assessment, and accountability systems.

Deputy Commissioner Leather joined a small planning taskforce supported by the Hewlett Foundation made up of national partners including the CCSSO, Carmel Martin of the Center for American Progress, and others, which resulted in a clear roadmap for an Accountability 3.0 model for a *51st State*, resulting in a white paper by principal authors Gene Wilhoit, Linda Darling-Hammond, and Linda Pittenger, (see Appendix D). This model includes components such as multiple indicators for academics, skills, and dispositions, an integrated local and state system of assessments, an inspectorate model of quality performance reviews of schools and districts, intense professional development, and comprehensive support for schools that

demonstrate performance issues. The PACE model was being designed concurrently with the 51st State model work with ongoing interactions among the two sets of designers to benefit from the thinking of each group.

Most importantly, Commissioner Barry has met monthly with district and school leaders ever since she came into office, engaging in deep conversations about how school accountability can best be designed to support significant improvements in school and student learning. These conversations led to the initial proposal to the USED in July, 2014, including two pilots, one based on the PACE work, the other on a College Readiness System connected to the suite of College Board Assessments. The NH Legislature, the Governor's Office, and other key stakeholders, such as the NH Institute of Higher Education Network, the School Administrators Association, the NH School Principals Association, the NH Chapters of both the National Education Association and the American Federation of Teachers are all in support of this new, more fully balanced system of reciprocal accountability based on the core principle of shared responsibility among state and local leaders.

Rationale for PACE

New Hampshire is committed to raising the bar for all students by defining college and career-readiness to encompass the knowledge, skills, and work-study practices that students need for post-secondary success including deeper learning skills such as critical thinking, problem-solving, persistence, communication, collaboration, academic mindset, and learning to learn. However, NH's educational leaders recognize that the level of improvement required cannot occur with the same type of externally-oriented accountability model that has been employed for the past 12 years. In fact, the state argues that the current system is likely an impediment for moving from good to great. The state intends to move to a model of an accountability system with significantly greater levels of local design and agency to facilitate transformational change in performance. As part of this shift in orientation, the state believes there are more effective ways to assess student learning for informing and improving students' progress. The State argues that a competency-based approach to instruction, learning, and assessment is philosophically and conceptually related to this internally-oriented approach to accountability and can best support the goal of significant improvements in college and career readiness. The information learned through competency-based assessments would then be used to make better accountability determinations that would better inform school improvement.

A competency-based system relies on a well-articulated set of learning targets that helps connect content standards and critical skills leading to domain proficiency. Such a system requires careful tracking of student progress and ensures that students have mastered key content and skills before moving to the next logical set of knowledge and skills. Current systems that rely on compensatory systems (e.g. averaging) for grading and related record-keeping may allow students to slip through the cracks in terms of possessing necessary knowledge for building deep understandings in the focal disciplines. Participating pilot districts will develop organizational and student-based goals and have these approved by a committee of peers as well as the SEA to ensure appropriate levels of rigor and focus on the key sets of knowledge and skills in each domain. Their progress towards meeting these defined goals will be reported by each participating district to the NH DOE at least yearly.

In the PACE option, the NHDOE has created a route for districts and schools to demonstrate progress that is not solely or primarily dependent upon state standardized tests. The creation of the PACE accountability option reflects NHDOE's belief that school accountability works best if the responsibility for design and implementation is shared by districts and the state, rather than top-down mandates. Known as "reciprocal accountability," districts and schools are responsible for determining and reporting on local accountability measures, while the state is responsible for support and oversight in helping districts establish strong accountability systems.

The PACE system is designed to foster deeper learning on the parts of students than is capable under current systems. Further, while NHDOE is a strong supporter and governing member of the Smarter Balanced Assessment Consortium, (SBAC), it argues that once/year assessments, as good as Smarter Balanced may turn out to be, is not enough to drive and support deeper learning. To do so requires timely assessments linked closely with curriculum and instruction. The PACE system is based on the belief that a rich system of local and common (across multiple districts) performance-based assessments is necessary for supporting deeper learning as well as allowing students to demonstrate their competency through multiple performance assessment measures in a variety of contexts. Thus, the accountability option was established to enable schools and districts to demonstrate student achievement and learning growth through means other than or in addition to standardized tests, with an emphasis on performance assessment.

Finally, New Hampshire is committed to implementing a philosophically coherent system. If the State is encouraging districts to embrace student agency in determining learning goals, then it only makes sense for the State to embrace "district agency" in establishing its own accountability goals.

Implementation Plan

NHDOE is engaged in a multi-faceted implementation plan to ensure the success of the PACE option that includes the following components:

- ✓ Requirements for participating districts
- ✓ Technical and professional learning support, including task development and scorer calibration activities
- ✓ Benchmarks for success
- ✓ Communication plan

Requirements for participating districts ("guardrails")

Districts participating in the 2014-2015 pilot must have already adopted the State graduation competencies and developed a coherent and high quality set of K-12 course and grade competencies mapped to the State graduation competencies. These districts must have demonstrated the leadership and educator capacity to participate effectively in the pilot. In addition to having a well-articulated set of competencies, these districts must have developed or be close to completing the development of a comprehensive assessment system. Because districts need to have demonstrated strong levels of performance and capacity in order to effectively participate in the pilot, districts with priority schools are not eligible to participate at this time. Districts considered for the 2015-2016 pilot must have adopted graduation

competencies and have a commitment during 2014-2015 to fully build out their course and grade competency systems in K-12 as well as their comprehensive assessment systems.

In order to be selected for the pilot, districts must be willing to participate in a peer and expert review process where they submit their system of performance-based assessments for evaluation based on clear and rigorous criteria including alignment with state standards and competencies, consistency and accuracy of scoring, and fairness to all test takers. Further, PACE districts will be required to administer the state summative assessments (Smarter Balanced) in at least three grades, one at each level (e.g., 4, 8, and 11), which will serve as both an internal and external audit regarding school and district performance. Local districts will be expected to incorporate the results of the Smarter Balanced Assessments in their local accountability system.

All pilot districts are expected to fully participate in the development and implementation of the pilot accountability requirements such that all pilot districts will have the same general assessment requirements in the same courses and grades. These general requirements are outlined out below in Table 1. As can be seen, the Smarter Balanced summative assessment will be administered in select grades. The current plan involves staggering the Smarter Balanced subject areas according to when the results will be most useful for informing programs and auditing the local and common performance assessments. The current state science assessment (NECAP) will be phased out as these districts play a lead role in beginning to pilot “next generation” science assessment tasks. In fact, the National Research Council advocated in a recent report¹ that moving to assessments of the Next Generation Science Standards must be led by classroom-based assessments rather than trying this extensive endeavor with large-scale assessments first. The PACE districts will be particularly suited to pilot this new approach, given their intensive efforts in implementing complex performance assessment both within and across disciplines.

Importantly, local performance assessment, used for competency determinations, will be administered in all subjects and grades. In certain grades and subjects, they will be “anchored” by Smarter Balanced assessment results, but in many others, they will be tied to the common performance assessments.

School districts participating in the PACE pilot will be required to report the number and percentage of students at each grade level who are meeting both locally defined, but state (and peer) approved definitions of proficiency and competency. Student growth for PACE districts will be reported according to locally determined and peer approved approaches for documenting student progress towards graduation competencies.

¹ National Research Council. (2014). *Developing Assessments for the Next Generation Science Standards*. Committee on Developing Assessments of Science Proficiency in K-12. Washington, DC: The National Academies Press.

Grade	Competency Grading	ELA	Math	Science
K-2	<input checked="" type="checkbox"/>	Local Performance Assessments	Local Performance Assessments	Local Performance Assessments
3	<input checked="" type="checkbox"/>	Smarter Balanced & Local PBAs	Common and Local PBAs	Local Performance Assessments
4	<input checked="" type="checkbox"/>	Common and Local PBAs	Smarter Balanced & Local PBAs	Local Performance Assessments
5	<input checked="" type="checkbox"/>	Common and Local PBAs	Common and Local PBAs	Common "Next Gen" PBAs
6	<input checked="" type="checkbox"/>	Smarter Balanced & Local PBAs	Common and Local PBAs	Local Performance Assessments
7	<input checked="" type="checkbox"/>	Common and Local PBAs	Smarter Balanced & Local PBAs	Common "Next Gen" PBAs
8	<input checked="" type="checkbox"/>	Common and Local PBAs	Common and Local PBAs	Local Performance Assessments
9	<input checked="" type="checkbox"/>	Interim Smarter Balanced & Local PBAs	Common and Local PBAs	Local Performance Assessments
10	<input checked="" type="checkbox"/>	Common and Local PBAs	Interim Smarter Balanced & Local PBAs	Common "Next Gen" PBAs
11/12	<input checked="" type="checkbox"/>	Smarter Balanced (plus optional SAT)	Smarter Balanced (plus optional SAT)	Local Performance Assessments
		*CAPSTONE PERFORMANCE ASSESSMENT		

Table 1. General assessment requirements for the PACE pilot accountability system.

The Task Bank

An ultimate goal of the PACE pilot is to enhance the capacity of educators to develop and use their own classroom assessments. However, creating a set of tasks for common administration and scoring purposes as well as helping to jumpstart local capacity is critical to the success of this project. The NH Task Bank is a repository of quality performance tasks that have been designed specifically to assess student attainment of the New Hampshire Competencies. In addition to serving as assessments for that purpose, the tasks in the NH Task Bank serve as models that teachers can refer to in their own assessment design work. We describe the current state of the NH Task Bank as well as the processes for continuing the development of the bank to support the success of the PACE system.

One of two key sources for performance tasks are those designed and submitted by New Hampshire teachers, most of who have participated in New Hampshire’s Quality Performance NH PACE Accountability Proposal

Assessment Cohort over the past three years. These teachers received training in task design, quality assurance, analysis of student work and calibration.

Tasks that are submitted to the NH Task Bank undergo a rigorous vetting process. A task is assigned to two reviewers. Typically, one reviewer is a Quality Performance Assessment Associate at the Center for Collaborative Education. The second reviewer is either a) a New Hampshire teacher who has worked closely with the Performance Assessment Cohort and become a peer coach within that Cohort, b) a Senior Associate from the Center for Assessment, or c) a member of the NHDOE. The two reviewers will use the Quality Criteria Review Checklist (found at <https://ccebos.box.com/s/3so0ipb5it5nocas83q1>), to vet the task and draw up recommendations for revision. The reviewers then either execute the revisions or return the task to the designer for revisions according to the recommendations. Once the task revisions are made and the task is accepted, it is posted to the Task Bank as “Accepted, Ready for Field Testing.”

The second key source of performance tasks is through the CCSSO’s [Innovation Lab Network](#) (ILN) Performance Assessment Project ILN, which began in the 2013, led by the [Stanford Center for Assessment, Learning, and Equity](#) (SCALE), the [Stanford Center for Opportunity Policy in Education](#) (SCOPE), the [Educational Policy Improvement Center](#) (EPIC) and the [Center for Collaborative Education](#) (CCE). The ILN project is collecting and curating a set of quality performance tasks that will populate an open-source, vetted task bank accessible to teachers. The emphasis of the work is on the type of performance-based measures that support assessment of the deeper learning.

The ILN Tasks were first pilot tested in spring 2014 and included 27 tasks in secondary ELA, mathematics, and science and was conducted in the following five states: California, Connecticut, Maine, New Hampshire, and Oregon. The tasks are aligned to the CCS standards for ELA and math, and the NGS standards for science. The next pilot stage of the performance tasks is scheduled to take place in spring 2015.

Similarly, tasks submitted for inclusion in the NH Task Bank are reviewed to ensure their consistency with PACE Quality Criteria and formatted to fit a common project template. As the spring 2014 tasks are revised based on the field test data they will also go through a process checking alignment with the New Hampshire State Competencies, and be placed in the Task Bank. NH tasks that were not included in the spring 2014 field test are being reviewed in the current round, ensuring a second level of quality assurance. A two way pipeline between the NH and national ILN task bank is envisioned.

The NH task bank currently includes more than two dozen tasks, that number is regularly increasing as tasks continue to be reviewed and added. Additionally, the ILN pilot task bank contains 27 tasks. A complete list of tasks with descriptions can be found at <http://tinyurl.com/alltaskslist>. The majority of tasks in the Task Bank are for the high school level, reflecting the initial focus on providing models for assessing New Hampshire’s graduation competencies. There is also good representation of tasks for elementary levels, which is why the current efforts are focused on developing common tasks for the middle school grades. The Task Bank contains tasks for which there exist approved New Hampshire Competencies: ELA, mathematics, science, and work-study practices (e.g., cross-cutting skills and dispositions).

Plans for expanding the New Hampshire Task Bank are underway. As tasks are submitted from PACE districts and other schools in the current and past New Hampshire Performance Assessment Cohorts, they enter the vetting process in an ongoing, rolling system manner. In addition to the ILN field testing, NH tasks will be field tested through the PACE work and the NH Cohorts. The plan is for the tasks to be revisited by summer 2015 when the student work generated by the tasks is available to be analyzed, for benchmarking and for further revisions to be made, if necessary.

Technical and professional learning support

The professional learning opportunities associated with PACE are couched in the actual work of PACE, including task development and scorer calibration activities. The implementing schools established work groups, creating common developmental competencies in the key content areas aligned to the state graduation competencies as well as continuing to build the state task bank. CCE is responsible for initial housing and maintenance of the state task bank and has now connected with the ILN-supported national task bank maintained by SCALE at Stanford University. In August of 2014, the PACE districts held two PACE Planning and Assessment Institutes, one a part of a larger state summer leadership institute and the other held at Sanborn Regional High School, in Kingston, NH.

In addition to ongoing reciprocal support and development work, the full set of PACE pilot technical advisors, along with district partners, will convene for a three all-day meeting to begin the formative peer review work. The dates and topics for these meetings are seen in Table 2 below.

Dates	Topics	Goals
Jan. 2015	Alignment & Fairness	<ul style="list-style-type: none"> • Build shared understanding of the expectations for documenting systems of assessments for alignment to learning targets and fairness • Share practice around assessment design and supports for all students
Feb. 2015	Monitoring Scoring (Calibration)	<ul style="list-style-type: none"> • Build a shared understanding of the expectations for scoring moderation (including calibration and collecting student work samples) • Share practice and lessons to date from district and cross-district content area working group calibration processes
May 2015	Fair & Comparable Competency Determinations	<ul style="list-style-type: none"> • Build a shared understanding of the expectations for aggregating evidence to make competency determinations • Share practice and identify next generation models for fair, mastery-based, personalized competency determination

Table 2. Major technical assistance convenings and formative peer reviews

We note that the technical review meeting schedule shown above is tied closely to the development and administration cycle. We first ensure that the tasks are aligned and fair, prior

to the administration of the common tasks. The next critical criterion is to document that raters can score consistently and accurately. Finally, once data are available, we will hold a technical assistance/review session to ensuring that competency determinations are created fairly, validly, and comparably.

In addition to these major technical assistance meetings, PACE district leaders and teachers, along with technical assistance providers will meet much more regularly to fully build out the performance assessment system for 2014-2015 (as seen below).

ELA PACE PLAN: DEVELOPMENT OF PERFORMANCE ASSESSMENTS				
GROUP	DATE & LOCATION	TASK	OUTCOME	NEXT STEPS
INTRODUCTION TO PURPOSE OF PACE ASSESSMENT FOR ELA				
SUMMIT DAY 1 (PACE DISTRICTS: ELA LEADS)	OCT 10 @SANBORN 9:00-2:00	<ul style="list-style-type: none"> Review grade level performance tasks for ELA specifications 	<ul style="list-style-type: none"> Create a plan 	Work with local grade level teachers.
AT HOME DISTRICT	OCT/NOV	<ul style="list-style-type: none"> Review Assessments Use the validation protocol to review the tasks Discuss the possible end of year ELA performance tasks 	<ul style="list-style-type: none"> Select 1 to 2 possible performance tasks for each grade level to share with the PACE group 	<ul style="list-style-type: none"> Prepare copies for ELA PACE Performance Task Summit to share with other grade level leads
SUMMIT DAY 2 (PACE DISTRICTS: ELA TEACHERS PER PACE GRADES)	DEC 5 @SANBORN 9:00-2:00	<ul style="list-style-type: none"> Review grade level performance tasks for ELA specifications Review shared materials and rubrics 	<ul style="list-style-type: none"> Create a plan Distribute shared rubrics 	<ul style="list-style-type: none"> Work with local grade level teachers. Align instruction to DOK 3 to prepare students for tasks
SUMMIT DAY 3 (PACE DISTRICTS: ELA LEADS)	FEB 13 @SANBORN 9:00-2:00	<ul style="list-style-type: none"> Discuss the materials for the implementation in March Finalize tasks/assessments 	<ul style="list-style-type: none"> ELA Leads to upload tasks/assessments to the task bank 	<ul style="list-style-type: none"> Prepare finalized tasks for teachers
ADMINISTRATION OF PERFORMANCE TASK/ASSESSMENT				
AT HOME DISTRICT	MAR (MONTH)	<ul style="list-style-type: none"> Administer the Performance Tasks/Assessments 	<ul style="list-style-type: none"> Review student work Discuss assessment administration List any changes that need to be made 	<ul style="list-style-type: none"> Compare PBA results with other assessments measuring similar targets (e.g., competency grades, NWEA/STAR/SBAC)
REFLECTION & DEBRIEF				
SUMMIT DAY 4 (PACE DISTRICTS: ELA LEADS)	APRIL 17 @SANBORN 9:00-2:00	<ul style="list-style-type: none"> Discuss changes to upload to the task bank Identify successes and challenges 	Create a plan for 2015-16	

Table 3. PACE ELA Task Development Plan for 2014-2015 (note: mathematics and science follow a similar plan).

Benchmarks for success

The activities described in the implementation section essentially constitute the benchmarks for success during this pilot year. We summarize these below, along with the reporting schedule.

Dates	Systems Completed	Specific Deliverables
12/30/14	Summative Performance Assessment Tasks	<ul style="list-style-type: none"> • Report describing status of common summative performance assessments to be used in 2014-2015 pilot by the four districts and uploaded to the State Performance Bank in the following content areas, (including work study practices embedded): <ul style="list-style-type: none"> ○ English Language Arts, ○ Mathematics, and ○ Science • Documentation of shared practice around assessment design for all local summative assessments that contribute to students' summative determinations, including alignment, rigor, and fairness
1/30/15	Scoring Quality of Performance Assessments	<ul style="list-style-type: none"> • Report describing status completion of: <ul style="list-style-type: none"> ○ common scorer moderation/calibration protocols developed and shared among all districts ○ Designing and using formative assessment processes to maximize student learning
2/28/15	Professional Development and Technical Peer Reviews	<ul style="list-style-type: none"> • Report on professional development activities within and across districts related to: <ul style="list-style-type: none"> ○ shared practice and lessons to date from district and cross-district content area working group calibration processes ○ to address a shared understanding of the expectations for aggregating evidence to make competency determinations from performance assessment results ○ regarding shared practice and identification of next generation models for fair, mastery-based, personalized competency determination ○ protocols and plans for the technical peer reviews of the four implementing districts
3/30/15	Competency Determination	<ul style="list-style-type: none"> • Report describing: <ul style="list-style-type: none"> ○ Specific plans for making annual determinations in all grades and subjects (ELA, math, and science)
4/30/15	Summative Assessment	<ul style="list-style-type: none"> • Report describing: <ul style="list-style-type: none"> ○ Status update on PACE Summative Performance Assessment and Smarter Balanced Window.

5/30/15	Summative Assessment	<ul style="list-style-type: none"> • Report describing: <ul style="list-style-type: none"> ○ Status update on PACE Summative Performance Assessment and Smarter Balanced Window, as per Table 1.
6/30/15	Summative Assessment	<ul style="list-style-type: none"> • Report describing: <ul style="list-style-type: none"> ○ Status update on PACE Summative Performance Assessment and Smarter Balanced Window, as per Table 1.
6/30/15	Increased and Sustained Professional Development	<ul style="list-style-type: none"> • NH Performance Assessment “summer camp” to both provide intensive training to build local expertise and to develop assessments to fill holes in the Task Bank.
7/30/15	Technical Peer Review	<ul style="list-style-type: none"> • The results of the complete first year technical peer review will be conducted and reported.

Communication

A major advantage of doing this work in New Hampshire is that it is small enough to enable a high degree of personal communication—everybody knows everybody! The state has only a small handful of daily newspapers and a limited number of other major media outlets include one television station and NH Public Radio, which does a fair amount of education reporting. Therefore, it is relatively easy to implement communication approaches in NH to ensure that all relevant stakeholders are informed and engaged in this and other initiatives.

First, the Commissioner and her senior staff meet with all of the district superintendents every month to present ideas and solicit feedback. Second, NH has had an active accountability task force since before NCLB was implemented. This advisory group, comprised of a broad-based set of practitioners, has been involved in designing, reviewing and commenting on multiple iterations of this proposal. Third, the state has been meeting with a policy oversight group to support implementation of the Quality Performance Assessment Initiative and the PACE accountability initiative.

NHDOE, with the support from the National Governors Association and the Helmsley Foundation, in conjunction with the Achieve Competency Education Network initiated a communications campaign led by NH State Board of Education member, Bill Duncan, through his highly respected and viewed website, Advancing NH Public Education (www.ANHPE.org). Additionally, the NHDOE, with 2Revolutions, LLC, launched the *The NH Story of Transformation* (<http://nhtransformation.squarespace.com>), which articulates the work completed over the last 20 years in developing competency education in NH leading to the new system of PACE accountability.

With the assistance of 2Revolutions and Advancing NH Public Education, NHDOE will be launching a PACE website to support internal collaboration and external communication. This site will provide the background necessary to understand the changes, where teachers, administrators and parents speak from their direct experience and students' display and discuss (in written and video form) their projects and their learning experiences.

Technical Issues and Considerations

The NHDOE has provided extensive documentation in previous meetings with USED regarding our plans for ensuring the technical quality of the PACE system. We review some of the key issues below, including:

- ✓ Comparability of annual determinations (status)
- ✓ Progress (growth)
- ✓ Equity
- ✓ Research and evaluation
- ✓ Utility

Comparability of annual determinations

One of the major challenges with the PACE pilot accountability system is ensuring that students from all NH receive meaningful opportunities to learn the required knowledge and skills. One of the ways to evaluate these opportunities is to require all students to participate in the same assessment of the required knowledge and skills. But it is not the only way. There are many examples, both with educational programs and outside of education, where we recognize that the “same” is not the only way to define comparability. For example, consider students applying for a competitive music program. Students will play different songs, perhaps using different instruments, but judges will have to determine who should be admitted to the program. We accept that judges are able to weigh the different types of evidence to make “comparable judgments.” Why do we accept this? Because we have great trust in expert judges and their shared criteria. When the criteria are not explicit and applied systematically, then people have concerns (remember some of the Olympic figure skating fiascos in past years).

We know that true psychometric comparability (i.e., “interchangeability”) across districts administering different systems of assessment cannot be assured. In fact, we know that it is not expected. However, NHDOE is taking important steps to ensure that students in pilot districts receive a high-quality education that meets or exceeds the expectations for non-pilot districts held to the same high expectations. For example, students deemed proficient in a particular grade or content area likely should be considered proficient regardless of the type of assessment.

We argue that comparability efforts should not be focused on individual assessments administered throughout the year, rather the focus of comparability must be on the annual determinations. NHDOE has proposed an approach to do just that. The Smarter Balanced Achievement Level Descriptors (ALDs) are the basis for establishing cut scores on the Smarter Balanced Assessments (this process was recently completed). The ALDs serve as the narrative descriptions of performance and the role of the standard setting panelists is to match the narrative descriptions with actual performance on the test. Therefore, NHDOE has decided to require all of the PACE districts to anchor their annual determinations of proficiency (competency) to the Smarter Balanced ALDs for the respective grade level and subject area.

Of course, it is one thing to use common descriptors, but having assessment evidence to evaluate against these descriptors is another critical component of comparability. Therefore, all PACE

districts have agreed to participate in a common standard setting process based on thoughtfully-identified set of summative competency assessments administered throughout the year along with the common summative PACE performance assessment. Participating in a common standard setting process, where student work is compared with the ALDs will allow for comparably rigorous achievement standards to be established in all PACE districts.

The use of Smarter Balanced ALDs and the common standard setting activity is the crux of the comparability approach, but this requires assurances that assessments are aligned with a common set of content expectations. To ensure that all students are held to the same set of college and career ready expectations, the state has adopted college and career readiness standards and state model competencies that describe the knowledge, skills, and work study practices that all students are expected to master before they exit the K-12 system. All districts, regardless of the assessment system in use, must demonstrate the alignment of their systems with these standards and model competencies.

To audit the extent to which the intended comparability has been achieved, NH DOE will rely on the results of the Smarter Balanced assessments in math and ELA in at least three grades (e.g., 4, 8, and once in high school). Additionally, we are closely examining the SBAC interim assessments to replace current local benchmark assessments, in our on-going efforts to marry local and state accountability systems. These common state assessments provide both an internal and external audit for locally-designed systems of assessment, evaluating the degree to which student performance on the local performance assessment system relates to performance on the statewide assessments. All districts participating in the PACE pilot will be expected to participate in a peer review process during the first two years of implementation in order to examine their system design, assessment results, and to explain/justify any discrepancies between the performance assessment and consortium assessment results. Peer review will be structured to provide support and technical assistance to districts to ensure that local systems maintain high quality.

In addition to these statewide assurances, the NHDOE supports PACE districts to establish assessment commonality and/or comparability among the pilot cohort. NH DOE is currently working with pilot and other districts participating in Quality Performance Assessment (QPA) cohorts to create and validate common performance assessment tasks (with accompanying guidelines, tools, rubrics, student work anchors, and data reporting processes) in each of five core disciplines (mathematics, English language arts, social studies, science, arts) and at each of the three grade spans (K-5, 6-8, 9-12). Each task will be constructed so that it can be curriculum-embedded and administered in local districts, and will be made available in a statewide online bank of performance tasks. The state expects PACE districts to draw from these common tasks as part of the agreement to participate in the cohort, and will also provide common protocols for creating and validating other locally-designed performance tasks.

Lastly, NH DOE is taking steps to ensure scoring comparability by promoting reliable scoring of performance assessment tasks across classrooms, schools, and districts. NH DOE is sponsoring Professional Development Institutes, including summer and school-year Quality Performance Assessment institutes on assessment literacy, competencies and designs for teaching them (knowledge, skills, and work study practices), assessment task design and validation, scoring

calibration, and data analysis to track student progress and inform instruction. Regional task validation sessions have been conducted to assist districts in fine-tuning assessment tasks to ensure they measure target knowledge, skills, and work study practices. Regional calibration scoring sessions will be conducted to build inter-rater reliability and consistency in scoring across districts. These sessions serve as professional development for participants to lead task validation and calibration scoring sessions at the local level.

Equity

First, the state argues that the competency-based educational system at the foundation of this pilot is, by design, more equitable because educators focus on the learning needs of every student and do not allow any students to fall through the cracks. That said, the state will continue to aggressively monitor and report the performance of student groups as outlined in New Hampshire's currently approved ESEA waiver. In addition, districts participating in the PACE pilot will be subject to additional examination of student group performance through their required participation in a peer review process to evaluate aggregate and student group performance results during the first three years of implementation.

Student Progress and Educator Effectiveness

SLOs continue to be the main component of NH's educator evaluation system for all NH districts. This was the clear intention of the NH Task Force on Effective Teaching. The state believes that they can successfully document changes in student learning while supporting positive changes in local assessment and instruction. Pilot districts, because of the improvements in their assessment capacity, will be able to produce higher quality SLOs than most NH schools and districts. Therefore, we think the question should focus more on can pilot districts produce valid educator evaluation results and less on specific (and distal) approach for calculating current achievement conditioned on prior achievement.

NH has been using Student Growth Percentiles (SGP, see Betebenner, 2009) for school accountability purposes for many years and plans to support districts in incorporating aggregate SGP results into educator evaluations starting in the 2015-2016 school year. The NH Task Force on Effective Teaching recommended that SGPs not be required to be used for individual attribution, unless the district's specific evaluation plan requires such use. The Task Force recommended, and NHDOE agreed, that aggregate SGPs must be used at least as part of a "shared attribution" approach according to a district's (or school's) theory of improvement (e.g., grade-level or content area teams). This is an important distinction because it is our intention to employ a similar—but not exactly the same—model in the PACE schools. In other words, NH proposes to use Smarter Balanced at select grades to be able to calculate SGPs and use the results aggregated at the school level. These school-level results can be used to audit the individual SLO results and compare the "growth" of students in the pilot schools with other schools in the state.

Research and Development

NHDOE has received funding from the Hewlett Foundation to support a national technical advisory committee (TAC) that will be used to help solidify the evaluation plan. Further, NH DOE will be required to complete an end-of-year report by October 2015 for the Hewlett Foundation that will document the technical assistance and professional learning opportunities experienced by the pilot districts. A similar document will be provided for the Nellie Mae Foundation for their support of the technical assistance to the pilot districts.

Currently, professional development and technical consulting efforts are being well-documented and districts are documenting their efforts and successes at promoting assessment (and instruction) literacy. NH DOE views this year as a “proof-of-concept” before launching the slightly larger effort next year. Further, while SGP calculations will not count for accountability in 2015, we will use the aggregate SGP results to help shed light on how well the pilot districts are performing on college and career ready standards compared to academic peers from non-pilot districts. To be fair, we cannot expect a one-year miracle. The type of change NH DOE is supporting does not occur overnight and it would be unfair to draw conclusions based on distal outcomes (e.g., student achievement) before the reform has had a chance to take hold. Nevertheless, NH DOE will work with its TAC to develop and begin implementing a comprehensive evaluation and research plan for this pilot. It is the intention of Commissioner Barry to establish a third-party evaluation of the process, identifying a provider in 2015, with the evaluation extending into succeeding years.

Utility

Noted educational researcher and technical advisor to the PACE project, Henry Braun, stated that utility is the most important technical criterion by which we should judge the quality of accountability systems. Utility refers to the degree to which the policy/accountability system is able to support its intended aims. In the case of PACE, this would mean that the accountability system provides structure and information to help transform educator practices and ultimately student learning. While it is still relatively early in the process, we think this video and transcript of two teachers from Rochester School District found [here](#) exemplifies the types of transformation we intend to support on a large scale.

NHDOE looks forward to engaging in this pilot so that we can maximize the learning of all of our students and educators. We are glad to partner with USED and national partners on this exciting effort that can serve as a model for states and districts throughout the United States.

Appendices

Appendix A: NH DOE ESEA Flexibility Waiver Concept Paper

Appendix B: A Vision for Improved Education Accountability Systems: *Recommendations from the CCSSO Accountability Advisory Committee*

Appendix C: PACE Guide

Appendix D: Accountability for College and Career Readiness: Developing a New Paradigm

New Hampshire Department of Education ESEA Flexibility Waiver Concept Paper

Concept Paper

New Hampshire Department of Education ESEA Flexibility Waiver Concept Paper

Concept Paper

Introduction: A Shared Vision for Student Achievement

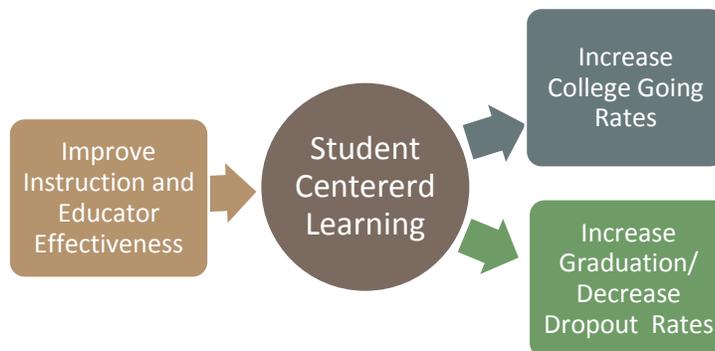
The New Hampshire Department of Education (NHDOE) is pleased to present this concept paper describing its vision for innovative approaches to improve student learning outcomes. The features contained in this paper will ultimately form the basis for a possible ESEA Flexibility Waiver application to the U.S. Department of Education.

Because of dedicated and focused leaders and teachers, New Hampshire has a long history of education excellence. This excellence is derived from a strong commitment to a shared vision for student achievement. This shared vision has allowed for the state to be regularly recognized for high student achievement, leadership and an overall quality education system. But as times change our strategies need to evolve. Now we must improve our system to ensure a better educational experience for all learners in a rapidly changing world – one that will result in more students reaching higher levels of learning, and better equipped to succeed beyond high school.

The NHDOE and the districts of the state will focus their efforts around four pillars of a high quality education system:

- Standards, Assessments and Instruction
- Data Collection and Use
- Teacher and Leader Effectiveness and Evaluation
- Transforming Struggling Schools

Through these pillars the state is committed to the following goals:



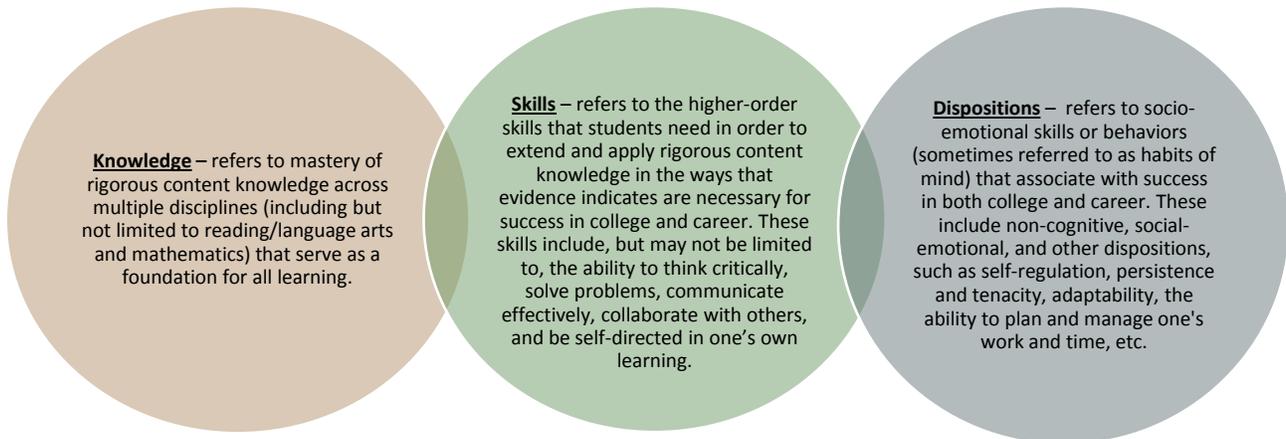
The NHDOE believes that for too many years, New Hampshire, along with every other state, has had to operate pursuant to an outdated federal education law that does not provide a rational accountability structure or the focused and meaningful supports our schools need. A request to the U.S. Department of Education to waive certain aspects of ESEA would allow for a more coherent overall approach thereby creating the opportunity for achieving dramatic improvements in student performance.

The state's current situation creates unnecessary complexity and confusion for New Hampshire schools, as evidenced by the current accountability structures being used. In 2009, the legislature of the State of New Hampshire passed SB180, a bill establishing an accountability system to ensure students receive the opportunity for an adequate education. The legislation specified a two-part accountability system: one part is input-based and the second part is performance-based. The input based component assesses whether a school provides the necessary curriculum for an adequate education and sets appropriate expectations for completion of the academic program. The performance-based component assesses adequacy based on the school's demonstration of student achievement, engagement and persistence to graduation. However, in addition to these two, the state is also required to comply with the federal accountability system that does not support either the input-based or proficiency-based components specified by state law. It is time to harmonize the federal and state systems into a single cohesive and meaningful approach.

This document is only a concept paper. Thus it contains only initial thoughts to drive further conversations and development. NHDOE hopes that these ideas will serve as a starting point for deeper engagement and conversations that will take place this summer. From this process will emerge a more specific set of strategies and plans that will guide our work and form the basis for a waiver application. By applying for the federal waiver, the NHDOE believes that the state will be able to move toward a system that is better for all students – a system that has a support orientation instead of a compliance orientation. While a number of the ideas contained in this paper could be implemented without a waiver, the added flexibility and relief from mandates that results from the receipt of a waiver will allow the state to more quickly reach its goal. In the end, it will be the students of New Hampshire that benefit from a better, more rigorous, more innovative, more meaningful education that prepares them for success in college and careers.

A New Theory of Action and Change

New Hampshire believes that all students must be college and career ready by the time they complete high school. This means not only meeting the content knowledge expectations of the Common Core State Standards (CCSS) in English Language Arts and Math, but also demonstrating necessary college- and career-ready knowledge, skills and dispositions¹. Our system must show that students are advancing not just by demonstrating growth in learning, but by demonstrating competency.



Competency education like this starts with a system of college- and career –ready standards implemented through a comprehensive and highly effective system of educator, school, and district supports. The system we envision will be based on the idea that all actors – teachers, leaders and the community – are engaged and share the intention and desire to help every student reach proficiency – a theory of positive intent. We reject the idea of a deficit model where schools and districts are identified as failing and “shaming by naming” is used as a method to increase student achievement. Ours will be a system where networks of educators and communities of learners will work on an “improvement-to-innovation” continuum to advance their practice and better support student learning while seeding transformation and the future of learning for New Hampshire students. *Figure 1* below illustrates this improvement to innovation continuum.

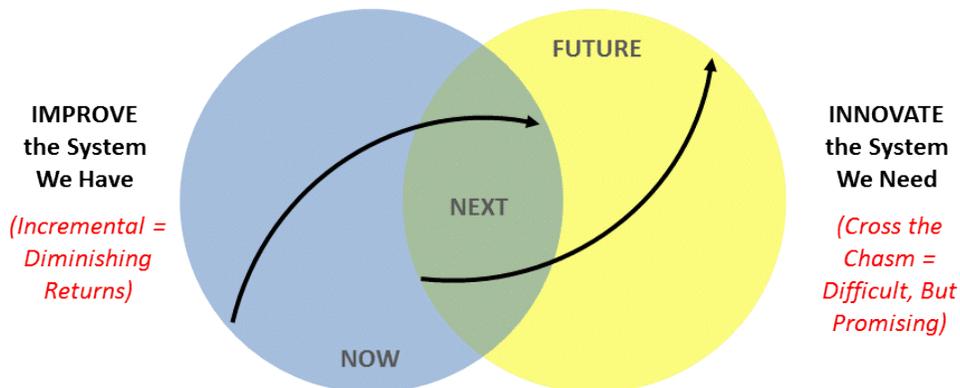


FIGURE 1- IMPROVEMENT TO INNOVATION CONTINUUM

Therefore, in New Hampshire:



- ▶ All graduating students will demonstrate college and/or career readiness based on an expanded definition of rigorous content and knowledge, adaptive skills, and critical dispositions by 2017. The state will also define ambitious but achievable annual measurable objectives (AMOs) that move beyond an accountability system based on a pure status model to one that eventually and fully includes a competency-based learning model.
- ▶ The state will adopt a balanced system of assessments (formative, interim, and summative) to assess student competency along learning progressions. Performance-based assessments will be administered when students are ready to demonstrate competency as opposed to waiting for an arbitrary date on a calendar.
- ▶ The state will set its ambitious annual yearly objectives (AMOs) with the intent of closing the gap of achievement in every subgroup by 50 percent by 2017 based on multiple measures.
- ▶ The state will provide a broad set of supports through a network strategy so that educators will be engaged in continuous, research-based improvement processes and support and cutting-edge, innovative approaches that rethink the structure of school practice and use of technology.
- ▶ The state will implement an educator effectiveness system connected to student performance, including competency attainment. It will address areas of preparation, selection, induction, mentoring and evaluation.

This new theory of change identifies areas of need and builds capacity for implementing the changes required to better meet the learning needs of all students, thereby becoming the newest chapter in the strong educational history of New Hampshire.

Principle 1: College and Career-Ready Expectations for All Students

NHDOE is committed to setting high expectations for what students must know and be able to do. NHDOE is a leading member of the Innovation Lab Network (ILN – coordinated by the Council of Chief State School Officers) and is primed to advance the work of the ILN around competency-based approaches to education. New Hampshire is prepared to provide a framework for schools to require their students to achieve proficiency of rigorous content knowledge and to demonstrate the ability to effectively integrate and apply knowledge in diverse environments within and across disciplines. NHDOE policy requiring all high school courses to be aligned to course-level competencies is one step the state has already taken to foster new practices of assessment that promote and evaluate deeper levels of understanding of important academic content, skills, and dispositions.

Adopting and Transitioning to College- and Career-Ready Standards

The New Hampshire State Board of Education adopted the Common Core State Standards (CCSS) in July 2010. Since its adoption, the state has provided multiple awareness meetings and professional development opportunities for educator engagement. To ensure consistency in messaging and to encourage a seamless transition, NHDOE developed an [implementation framework](#) (linked here) to help guide the work around the state, as well as efforts within districts and schools. The next step for the state will be to focus its implementation efforts on providing distinct and focused learning opportunities for the specific needs of district and school leaders, teachers, parents and students.

In addition to the foundation CCSS provides, New Hampshire has expanded the definition, or rather dimensions, of college and career readiness to include the knowledge, skills and dispositions students need to succeed beyond high school. Readiness requires more than students reaching higher levels of learning (as specified by the content standards).

New Hampshire’s dimensional elements of college and career readiness serve as a guidepost for dramatic reforms in education policy and practice. They represent the belief that we must deliver on the development of both cognitive and non-cognitive skills for all students as part of a moral, economic, and civic imperative to reduce inequities and advance excellence.ⁱⁱ

College and career ready means that students graduate from high school prepared to enter and succeed in postsecondary opportunities – whether college or career – without need for remediation.

- Students should graduate fully prepared to pursue the college and career *options of their choice*.
- *College ready* refers to the full range of programs leading to valuable, recognized degrees, including community colleges and four-year colleges.
- *Career ready* refers to employment opportunities with meaningful opportunities for advancement as well as career training programs that offer technical certification or other marketable skills.
- Evidence and experience indicate that the knowledge and skills needed to succeed in college and career are *greatly similar*, and that all graduates will need some form of postsecondary education or training to succeed during their careers.

To be college and career ready, students must graduate with the *knowledge, skills, and dispositions* necessary to succeed. These are the kinds of *deeper learning outcomes* that are at the heart of being college and career ready.

- Knowledge, skills and dispositions are mutually reinforcing, and not contradictory. That is, evidence and experience confirm that education that advances application of knowledge through skills is more likely to result in student competency of the underlying, rigorous content knowledge.
- The knowledge, skills and dispositions have concrete meaning and can be expressly taught, learned, and measured. This will require multiple, robust measures of evaluation and assessment.
- This same set of knowledge, skills and dispositions is also vital for student success in terms of citizenship, in addition to college and career readiness, including the ability to contribute and succeed in our increasingly diverse, democratic, global society.

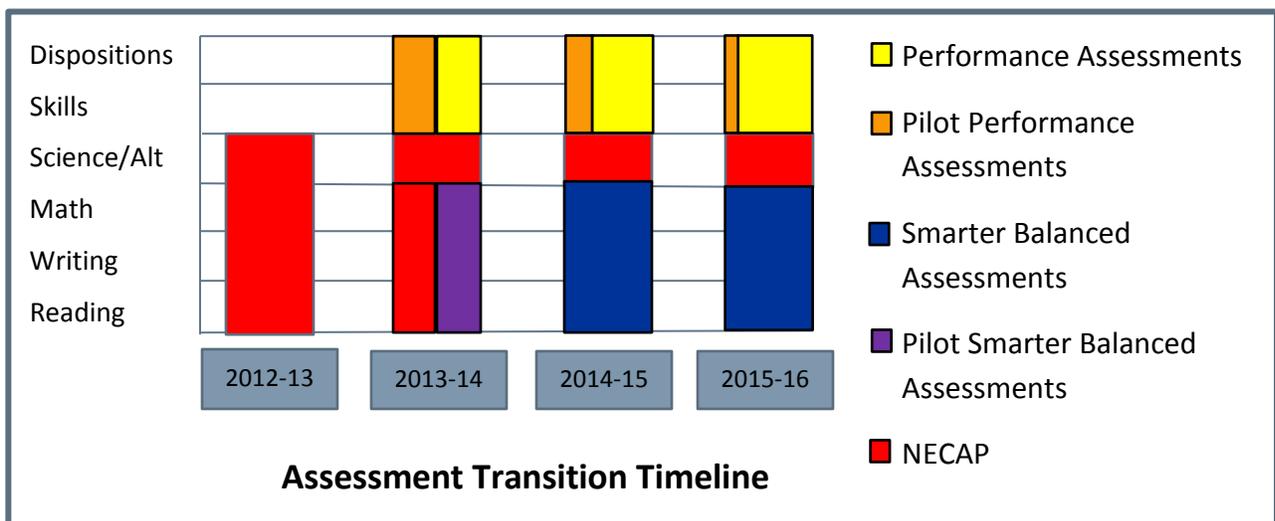
Developing and Administering Assessments that Measure Student Growth, Skills and Dispositions

In the context of the CCSS, and New Hampshire’s dimension of college and career readiness, the state needs ways to measure whether students are meeting expectations and reaching academic achievement goals. By 2015, the NHDOE is committed to creating a balanced and robust system of assessments (formative, interim and summative) focused on personalized learning that will evaluate students’ competencies over rigorous academic content, adaptive skills, and critical dispositions. One component of this system will be the assessments being developed by the Smarter Balanced Assessment Consortium (SBAC), for which New Hampshire is a governing state. The state is currently working with SBAC and partner states to investigate how the Smarter Balanced assessments might also be used in a competency-based instructional model and graduation system.

Another component of the state’s system will include performance assessments that will be designed in partnership with the Center for Collaborative Education and the National Center for Assessment. These assessments go beyond the assessment of academic content and will allow schools to evaluate a student’s readiness through deeper diagnostics of their skills and dispositions. The state will work with K-12 educators to develop a series of rubrics to identify competency definitions and levels for knowledge, skills and dispositions to assure comparability across school districts. Since these broad tasks can be demonstrated in numerous ways, the state will work with districts to create both common and unique assessment tasks that can be used by students, as well as guidance for students to create unique learning experiences that can be assessed using these rubrics.

Performance assessments are defined as multi-step, complex activities with clear criteria, expectations, and processes that enable students to **interact with meaningful content** and that measure how well a student **transfers knowledge** and **applies complex skills** to create or refine an original product and/or solution.

Finally, the state will continue to offer the New England Common Assessment Program (NECAP) assessment for both science and alternative assessments. The state’s assessment system will also balance local control with state-wide accountability and comparability. The chart below shows the expected timeline to develop and implement the assessment system. The SBAC and performance assessments will begin with pilot sites before going to full scale.



Principle 2: State Developed Differentiated Recognition, Accountability, and Support

When fully implemented, New Hampshire’s new assessment system will serve as a foundation for strong teacher, school and student accountability systems that will allow the state to realize its new Theory of Action. The accountability system will meet federal requirements and also help promote and incentivize continued improvement of instruction and assessment. The accountability system will be designed with the goal of moving away from branding schools through a negative and unproductive process, and moving toward a process of promoting improvement and innovation. It will move beyond a pure status model to one that includes measures of growth and eventually proficiency of learning. The rich performance tasks that will be developed in performance-based assessments can help support educator evaluation systems by providing a means of documenting student learning that is attributable to an individual teachers or groups of teachers. Similarly, student performance assessment results will be a more accurate key component of school accountability in New Hampshire and will serve as a tool to differentiate and disseminate recognition and support.

Implementing a Differentiated Accountability System with Ambitious and Achievable Annual Measurable Objectives (AMOs)

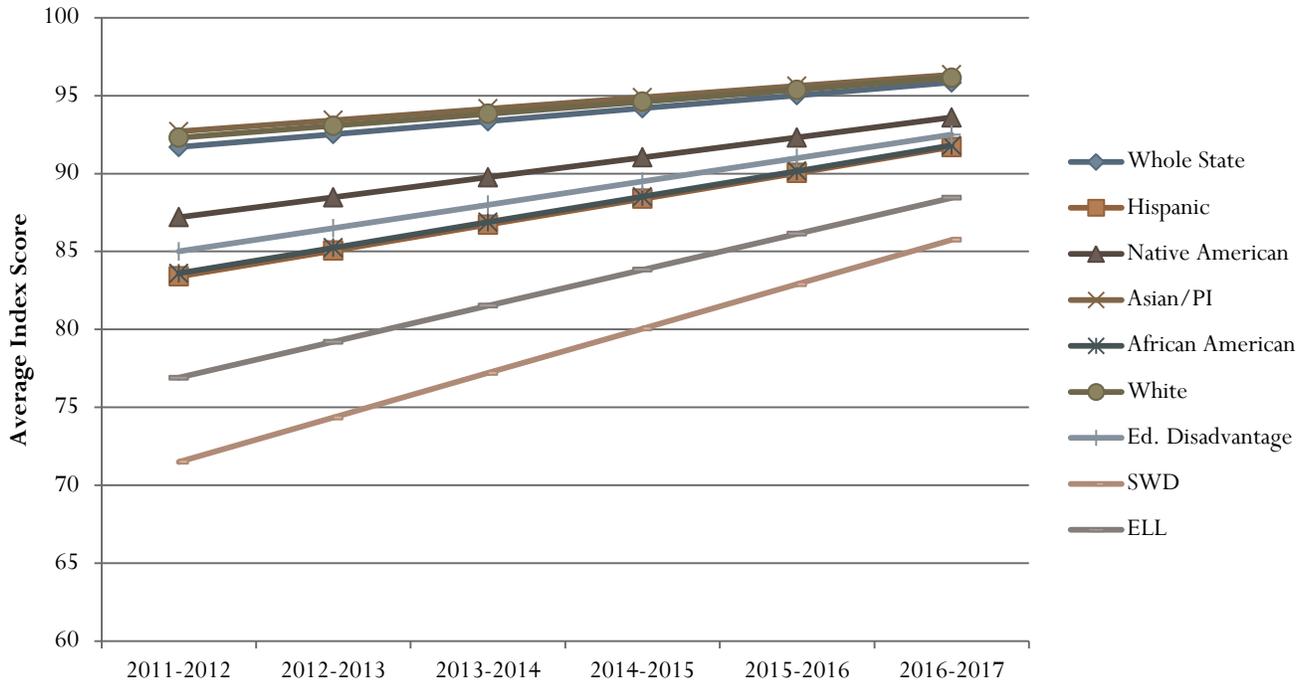
New Hampshire is committed to improving educational outcomes for **all** students, but recognizes that the current requirements of ESEA create obstacles to focusing on the schools and districts needing the greatest assistance. The state will take advantage of the opportunity afforded by USED to define and use more realistic AMOs. These will allow the state to differentiate levels of support for schools by building networks of technical assistance, knowledge sharing and innovation.

The federal flexibility waiver offers three options for redefining the states AMO determinations. New Hampshire will choose an option which allows the state to increase targets in annual equal increments toward a goal of reducing by half the percentage of students in the “all students” group and in each subgroup who are not proficient within six years.

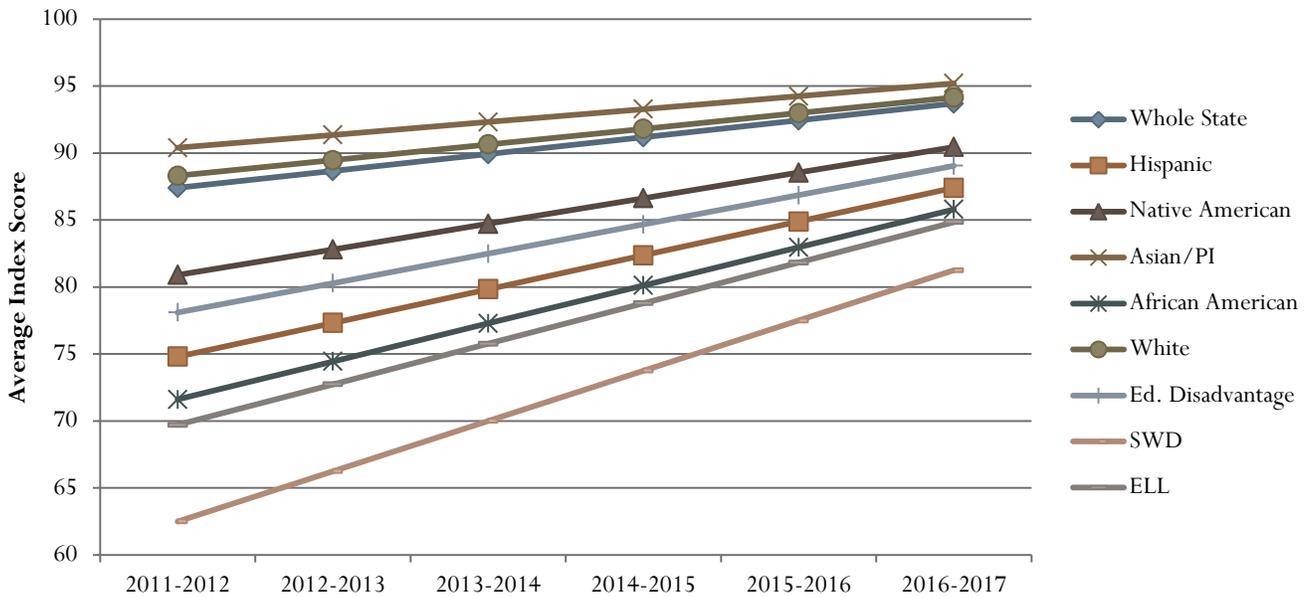
The state will initially define its AMOs based on the NECAP results now and for the next year. However, it is the state’s intention to reflect the multiple proficiency assessments it will develop – as well as the SBAC assessment – as those measures go into place. Throughout the summer of 2012, the state will work with stakeholder groups to determine the best way to include skills and dispositions into this new AMO determination.

The two charts on the following page demonstrate using the state’s current NECAP summative assessment performance by subgroup in Reading and Math.

**Principle 2, Option A: Elementary-Middle Schools AMO Calculations Based on
NECAP Reading Index Scores**



**Principle 2, Option A: Elementary-Middle Schools AMO Calculations Based on
NECAP Math Index Scores**



Identification and Support for Priority, Focus and Reward Schools

The NH DOE will continue to identify its lowest five percent *priority schools* and its additional 10 percent *focus schools* in the same manner as it has in the past few years. All available student achievement data for the past four years— using NECAP – for the “all students” group is reviewed for each school annually. The raw student achievement data for the state’s reading and mathematics assessments is converted to a 100-point index score. The index scores in each content area for the “all students” group are added together for each school in order to produce an annual combined score. The annual combined scores are then totaled to produce a cumulative achievement score for each school.

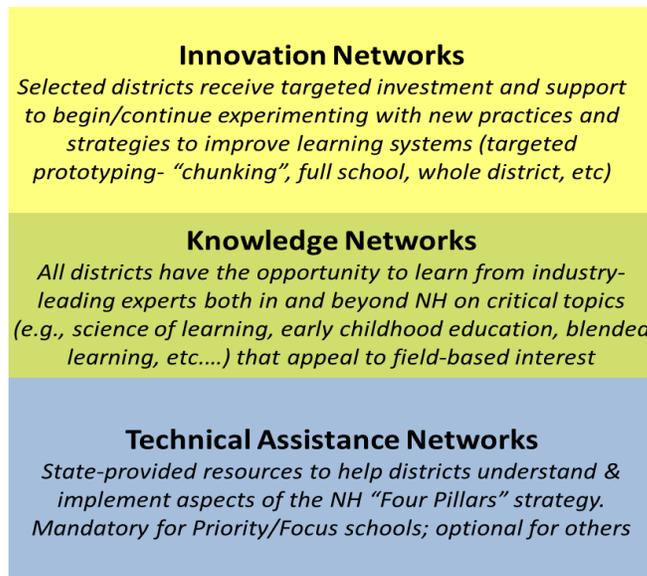
Schools are ranked in order from lowest to highest on the basis of the cumulative achievement score. Those at the top of the rank-ordered list are determined to be the state’s persistently lowest-achieving schools. The top five percent will be considered the state’s priority schools and the next 10 percent will be considered the state’s focus schools.

Priority and focus schools will receive intensive support and guidance from the NHDOE through its technical assistance networks (see next section). The details of this support, along with the determination and recognition of reward schools will be developed throughout the summer of 2012.

Developing Networks of Support and Recognition

New Hampshire is committed to implementing a new way of supporting its districts and schools. Shifting from a compliance orientation to a support orientation, NHDOE is reorganizing its structure, staffing and resources to better meet the needs of districts. In order to realize this shift, NHDOE is moving to a network structure of supports, aligned with the state’s Theory of Action but equally responsive to the needs and interests of districts and schools. This approach is based on the state’s recognition of a continuum from improvement to innovation, with the need to engage all districts and schools in the necessary work of continuous improvement while at the same time seeding the transformation of structures, practices and technology tools, which will yield models that are more personalized, rigorous and ultimately cost-effective.

The graphic below illustrates a comprehensive network strategy:



This network strategy will build capacity and provide supports for all schools and districts. Under this approach:

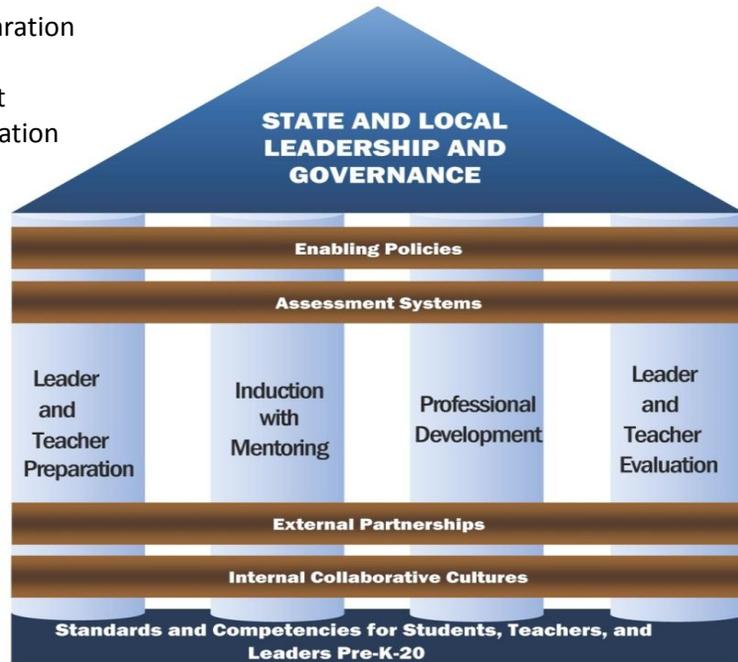
- **Technical Assistance Networks** will focus on continuous improvement. These networks will convene with regularity and be concentrated in 5 geographic regions of the state (North Country, Lakes, South West, South Central, Seacoast). Sample topics for these networks will include: Common Core Instructional Strategies, Common Core Implementation, Competency-based Grading & Assessment, Use of Data to Drive Continuous Improvement, Teacher/Leader Evaluation & Effectiveness, Literacy Across the Curriculum, Developing and Supporting Authentic Assessment, etc.
- **Knowledge Networks** will provide a range of stakeholders throughout the state with access to information based on needs and interests. These “networks” will be informal and will include tools such as blogs, listservs, webinars, conferences, seminars, and symposia, with a goal of encouraging intellectual discourse and attracting national and international thinkers to the Granite State. Work from both the Technical Assistance and Innovation networks can bubble up to the Knowledge Networks as a way to share emerging practice, new knowledge or dialogue about open questions. Sample topics for these networks include: Personalization Technologies, Early Childhood Education and the K12 System, Next Generation Learning, Cultural Shift to a Competency-based System, Developing Student Voice, Transdisciplinary Learning, Key Dispositions of Successful Students, etc.
- **Innovation Networks** are non-geographic opportunities for colleagues from around the state to come together around areas of shared interest with a problem-solving orientation in service of transforming the existing educational model by building alternative structures, practices and tools. These networks will be structured as limited time engagements with up-front facilitation by content & subject specific experts to develop ideas that can be rapidly prototyped in the classrooms, schools and districts. Sample topics may include: Mass Customization/Personalized Learning, Innovative Use of Time, Innovative Use of Space, Performance-based Data Management, New School Models, etc.

Principle 3: Supporting Effective Instruction and Leadership

The NHDOE – in partnership with educators and other stakeholders – are currently involved in developing a Comprehensive System of Educator Effectiveness characterized by four pillars:

- ▶ Leader and Teacher Preparation
- ▶ Induction with Mentoring
- ▶ Professional Development
- ▶ Leader and Teacher Evaluation

As the system has been conceptualized and constructed, state and local leaders have been diligently working over the last two years on developing and implementing policies, assessment systems, external partnerships (in-state, regional, and national), and internal collaborative cultures within schools and districts.



The NHDOE has undertaken a comprehensive process to develop model evaluation systems for both principals and teachers. This process has included over 100 stakeholders serving as thought partners on task forces and committees and is ongoing with teams meeting currently. The following provides a snapshot of the work conducted.

Principal Evaluation and Support Systems

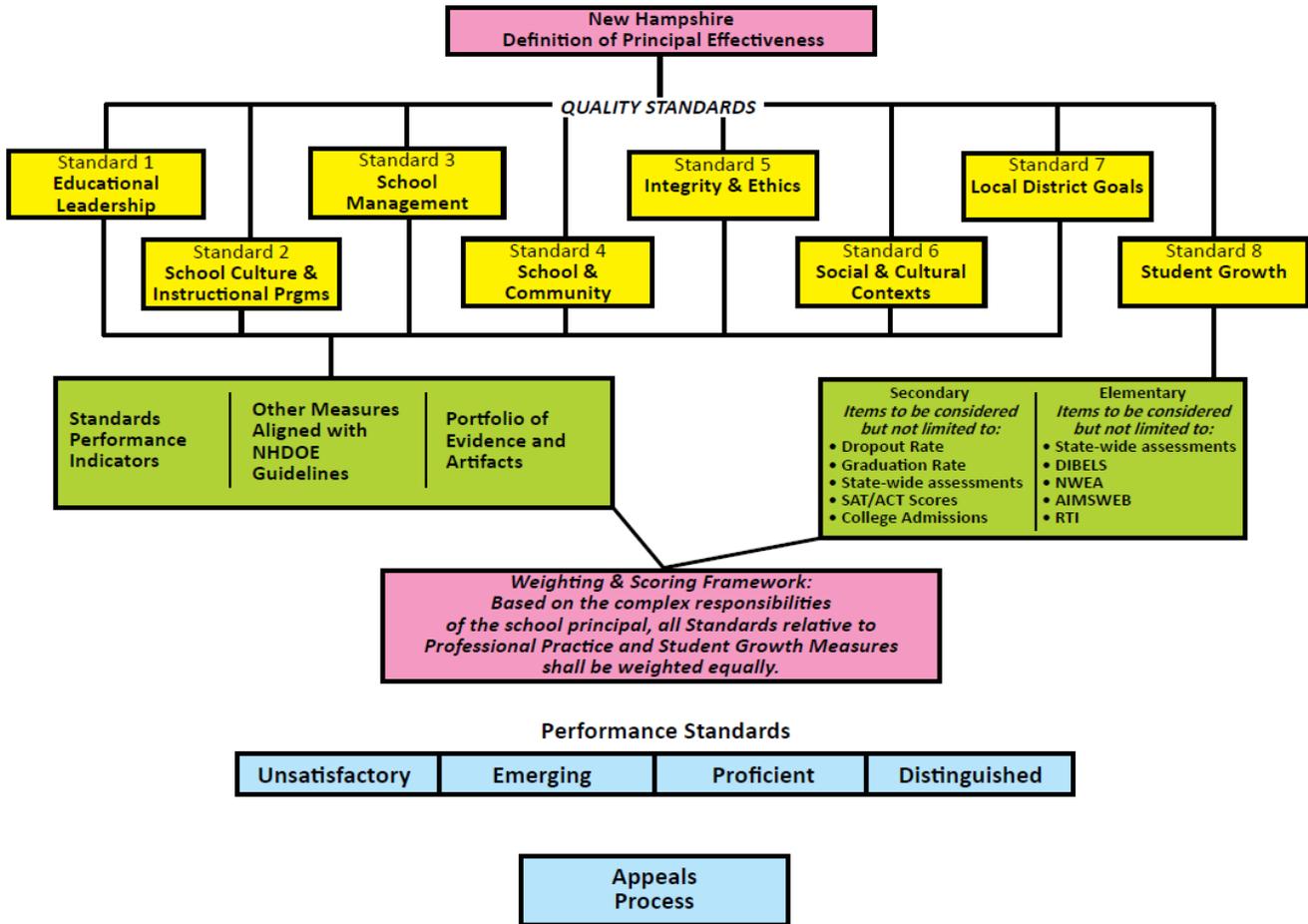
The New Hampshire Association of School Principals established a Principal Effectiveness and Evaluation Task Force to make recommendation on how principals should be evaluated. Their task was to:

- Provide a common definition of effective leadership at the principal level;
- Identify frameworks that are researched-based that might be used for fair and equitable evaluation processes; and to
- Develop a set of recommendations that will lead to supporting a framework for preparing, evaluating and supporting principals.

The Task Force defined effective principals as those who:

...promote the success of all students by facilitating the development, articulation, implementation and stewardship of a vision of learning that is shared and supported by the school community. An effective principal promotes the success of all students by advocating, nurturing and sustaining a school culture and instructional program conducive to student learning and staff professional growth. Principals are educational leaders who promote the success of all students by collaborating with all families and community members, responding to diverse community interests and needs, and mobilizing community resources.

The Principal Task Force recommended nine procedures and developed a framework for principal evaluations based on the ISSLC Standards.



Teacher Evaluation and Support Systems

Established in 2010 to build a foundation for the development of a system to support effective teaching in New Hampshire, the NH Task Force on Effective Teaching (Phase I) was comprised of sixty representatives from a wide range of stakeholder groups. The 2011 [Phase I Report](#) (linked here) contains details from this effort, including a common definition of effective teaching for all schools.

Effective teachers are those that focus relentlessly on the achievement of their learners. They are also deeply committed to the success of all learners. Research has shown that teacher’s knowledge and skills are in key areas – the learner and learning, content knowledge, instructional practice, professional responsibilities and dispositions – contribute, in varying degrees to student growth and achievement.

The Task Force also (please see [Phase I Report](#)):

1. Identified different teaching frameworks that are research-based and are critical components to a fair and equitable teaching evaluation process;
2. Developed a system of preparation, professional development, and continuous advancement of teachers to impact student learning; and
3. Developed a set of recommendations that will lead to a statewide system of teacher effectiveness.

The Phase II Task Force on Effective Teaching is currently active and is charged with carrying out the recommendations put forth in the *Phase I Report*. The Phase II Task Force is comprised of over 40 key education stakeholders, including teachers, principals, superintendents, higher education representatives, and key union and association representatives. The Task Force is supported by the NHDOE, the Center for Assessment, and the New England Comprehensive Center.

Guiding Principles

The primary purpose of the state model system is to maximize student learning and its development was guided by shared design principles. The following are highlights of some of the principles supporting this primary purpose.

1. High quality teachers are critical for fostering student learning. Therefore, the system is designed to maximize educator development by providing specific feedback that can be used to improve teaching quality.
2. Local evaluation systems must be designed collaboratively among teachers, leaders, and other key stakeholders such as parent and students as appropriate.
3. The state model system and all local systems will be comprehensive and, to the maximum extent possible, research-based.
4. The effectiveness rating of each educator will be based on multiple measures of teaching practice and student outcomes including using multiple years of data when available, especially for measures of student learning.
5. The model system is designed to be internally coherent and also compliment the NH Leader Evaluation System.
6. The educator evaluation system need to be seen as providing information for school principals and/or peer teams to ultimately make recommendations about each educator's effectiveness determination.

General Evaluation Framework

The state model system contains five major components, four domains of professional practice and one domain of student performance results. Each of the domains should be weighted relatively equally, although local districts have the discretion to adjust the weighting of the domains to reflect local priorities.

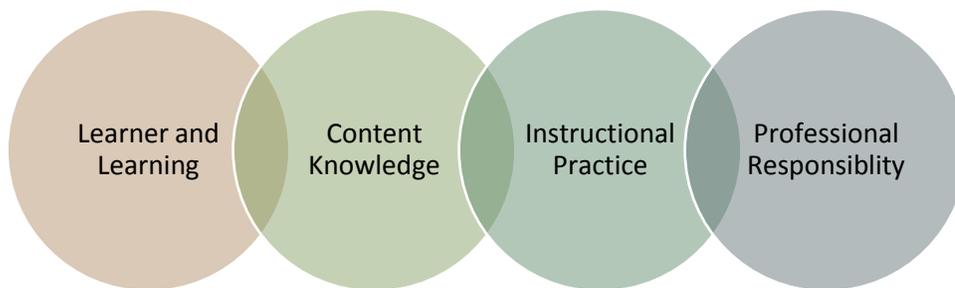
Each educator evaluation will include:

- Yearly self-reflection and goal setting
- A professional portfolio documenting key aspects of teacher practice
- Observations of practice by educational leaders and potentially peers
- Student Learning Objectives (SLO)
- Student Growth Percentiles (SGP -- if applicable)
- Shared attribution of at least part of the SLO and/or SGP results depending upon local theories of action around school improvement.

In addition to the major components listed above, *the Task Force recommends exploring the inclusion of measures of student voice and parent opinions in the evaluation of educators.*

Standards of Professional Practice

The Task Force recommends that all local systems should be based on the four domains of effective teaching described in the Phase I report:



The Task Force expects that districts will use other existing frameworks (e.g., Danielson) to help specify and measure the various aspects of professional practice, but recommends that all districts must map their framework to the four dimensions from the Phase I report.

Use of Student Performance Results

All teachers, whether in “tested grades and subjects” or not will be required to document student academic performance each year using Student Learning Objectives (SLO) in accordance with the SLO guidance developed by the Task Force. Student Learning Objectives is a general approach (also called Student Growth Objectives or Student Learning Evidence) whereby educators establish important goals for individual or groups of students (in conjunction with peers and administrators) and then evaluate the extent to which the goals have been achieved.

The NHDOE will produce Student Growth Percentiles (SGP) results documenting the individual student and aggregate growth for students based on state test data. These results will be aggregated according to “teacher of record” rules and for the whole school. Further, results will be disaggregated according to identifiable student groups in the school. These results, based on NECAP and eventually Smarter Balanced assessments, using the SGP model, will be incorporated into teachers’ evaluations either using

a shared or individual attribution framework. Both SGP and SLO analyses will produce results in three classifications of performance, to the extent possible, such as: high, typical/average, and low.

Coherence

The state model is designed to maximize coherence among the various aspects of the system. In particular, the Task Force wants to ensure that the four domains of teaching practice and student performance results are seen as integral parts of a comprehensive system. For example, this means that, to the extent possible, observations of teaching performance should be connected to measures of student performance (via SLOs) as a way to triangulate information. Similarly, the quality and usefulness of student performance measures should be incentivized and recognized as part of the specific domains of teaching practice.

Frequency of Evaluation

The frequency of summative evaluation will be tied to educators' length of time teaching and previous evaluation rating. Highly effective, experienced teachers will undergo a summative evaluation at least once every three years, while new and/or teachers previously rated ineffective will be evaluated every year. All teachers, however, will be expected to receive formative feedback and participate in SLOs and the professional portfolio process each year.

Consequences and Supports

The system has been designed to ensure that teachers with low evaluation ratings receive support in order to improve their teaching performance. If the performance of teachers on a continuing contract, as reflected in the evaluation scores, was low for a second year, the level of support should be intensified for at least another year.

At the other end of the continuum, teachers with exemplary performance as demonstrated by the evaluation ratings will be recognized in ways determined by the local district. This recognition may include monetary rewards, but more likely will include recognition and taking on additional responsibilities (e.g., mentoring, serving as evaluators) and perhaps additional flexibility from other requirements.

Implementation

Implementation will begin in the 2012-2013 academic year with volunteer districts and the School Improvement Grant (SIG) schools. The volunteer districts will be considered part of the first pilot phase. The second year of piloting will occur in 2013-2014 and will include the volunteer districts from the 2012-2013 as well as new volunteer districts. All districts will be expected to implement the state model system or locally aligned system by the 2014-2015 school year.

Principle 4: Reducing Duplication and Unnecessary Burden

If the NHDOE determines to move forward with the federal flexibility waiver, the state will use the process of designing the request and its current reorganization efforts to focus on reducing reporting and other burdens for districts. A stakeholder team will be brought together this summer to assist the NHDOE in determining the necessary and desired action steps for Principle 4. Ultimately, it is our goal that this process yields a more efficient, more effective organization.

ⁱ CCSSO Innovation Lab Network, College and Career Readiness Task Force. Draft *Framework for Defining College and Career Readiness*. May 2012

ⁱⁱ CCSSO Innovation Lab Network, College and Career Readiness Task Force. Draft *Framework for Defining College and Career Readiness*. May 2012

A Vision for Improved Education Accountability Systems

Recommendations from the CCSSO Accountability Advisory Committee

Draft – November 7, 2013

The Vision

The purpose of this document is to describe a vision for improved education accountability systems to inform policy and practice. To develop this vision, the Council of Chief State School Officers (CCSSO) convened a broad-based group of education leaders, the Accountability Advisory Committee, to develop recommendations that can be used as a resource informing the continued leadership by states and CCSSO as they advance systems of college and career ready accountability and supports (a list of members is provided as Appendix A). The advisory Committee considered current promising practices as well as new ideas that go beyond established systems. The recommendations of the Committee are summarized in this document.

The Committee was guided by the perspective that the time is right to revisit the established architecture of top-down, standardized, state accountability systems, based primarily on end of year summative achievement test scores. Moving forward, next generation accountability systems should:

- include a broader set of outcomes to more authentically reflect the breadth and depth of readiness for college, career, and citizenship;
- provide states and districts the autonomy to define some of the outcomes for which they will be held accountable;
- permit flexible testing approaches at the point of student readiness;
- be designed and evaluated for continuous improvement, with a clearly articulated link between design features and improved student achievement.

These principles form the foundation of the Committee’s recommendations. In the remainder of this vision document, a rationale supporting these recommendations will be provided along with specific recommendations for policy and practice.

Background

As states continue to lead efforts to transform our education systems so that all students graduate from high school ready for success in college and career, we must ensure that state systems of accountability and supports are best designed to support that transition.

In 2011, CCSSO released a set of principles (endorsed by 44 states and Washington, DC) to guide state leadership in raising the bar toward college- and career-ready (CCR) systems of accountability and supports, and called on Congress and the US Department of Education

(USED) to support this state leadership through the reauthorization of the Elementary and Secondary Education Act (ESEA) or ESEA waivers. Articulated in CCSSO's *Roadmap for Next-Generation State Accountability Principles*¹, the principles emphasized, among other elements, a focus on college and career readiness and a call for inclusion of differentiated performance indicators to inform support. In addition, CCSSO's principles committed all states to processes of innovation, evaluation, and continuous improvement in CCR accountability and supports. To further state leadership on the support systems as they relate to accountability, CCSSO (with support from EducationCounsel) established a Committee to deepen its work on several of the principles focused on diagnostic review and supports and interventions. The Committee released a resource paper earlier this year with additional guidance on using information from accountability determinations to inform effective supports and interventions.

In the absence of legislative action to reauthorize ESEA, in the fall of 2011 the USED responded to states leadership in developing a vision for accountability and launched an initiative to provide some flexibility from the NCLB accountability requirements aligned to that vision. Overwhelmingly, states applied for and were granted NCLB waivers. As of October 1, 2013, 42 states, the District of Columbia and Puerto Rico were approved. These waivers gave states an opportunity to develop and implement new systems of accountability aligned to CCSSO's principles, moving beyond NCLB to allow for more thoughtful and effective systems of accountability anchored in college and career readiness.

Subsequently, CCSSO (with support from the Center for Assessment and EducationCounsel) established this Committee on the future of accountability, with particular leadership from the CCSSO Innovation Lab Network (ILN) states. The work of this Committee was premised on the notion that this is a critical moment in state transition; states are in the midst of implementing new CCR systems of accountability and supports, states are transitioning to CCR assessments (aligned to Common Core or other CCR state standards) that raise the expectations for student outcomes, and underlying both of these, states are focusing on the kinds of shifts in teaching and learning that are needed to promote CCR outcomes for all students. To inform this work, the Committee was designed to take a step back – reflecting on the core goals and role of accountability – in order to then step forward toward the accountability systems states need going forward.

The summary of discussions and priority areas identified below broadly affirm CCSSO's prior accountability principles, build upon these principles, and seek to extend and improve systems in a manner consistent with the proposed vision. This has implications for states' continued leadership, as well as for ESEA waivers and waiver renewals.

¹ www.ccsso.org/Resources/Publications/Roadmap_for_Next-Generation_State_Accountability_Principles.html

Goals

Education accountability designs must be guided by the goals of the accountability system. These goal statements, which make explicit the intended outcomes, also serve as a foundation for the evaluation of the validity of the accountability system.

CCSSO's 2011 *Roadmap for Next Generation State Accountability Systems* suggested four goals for education accountability systems, which are summarized as follows:

- I. Articulate expectations in order to align actions and decisions toward ensuring all students are ready for college and career.
- II. Differentiate the performance of schools and districts in meaningful ways so that those in need of improvement receive appropriate supports and interventions, and those excelling can be recognized as models of excellence.
- III. Provide transparent, timely reporting of actionable data on performance results so that stakeholders at all levels can take appropriate action.
- IV. Foster a commitment to innovation and improvement.

The Committee upholds these goals, while extending them by offering the following areas of emphasis:

1. **Readiness:** A central goal of education systems is to improve student achievement such that all students graduate ready for college and career, but the Committee encourages states to broaden their view of CCR knowledge, skills, and dispositions. College, career, and civic readiness must encompass not only mastery of rigorous content knowledge but also the ability to apply knowledge through higher-order skills and underlying learner-ready dispositions. The Common Core state standards and other college and career readiness standards represent a “core” of the knowledge and skills needed for CCR, and can be implemented in a manner that has the biggest possible impact on student mastery of CCR knowledge, skills, and dispositions. Still, states should go beyond Common Core standards and assessments to more broadly define and measure the full range of desired CCR skills and dispositions.
2. **Academic Achievement and Growth:** Accountability systems should promote not only district and school performance improvement but also individual student achievement and growth, which can be supported through the provision of timely, actionable information by which stakeholders can make decisions regarding instructional practice.
3. **Equity:** Accountability systems should incentivize improved academic achievement for all students, including and especially students who are at risk. This focus on equity is not in conflict with efforts to promote gains for high performing students.
4. **Effectiveness:** Accountability system designs should encourage, and not impede, the kinds of shifts toward personalized or performance-based teaching and learning necessary to achieving CCR outcomes for all students. Considerations should also include the provision of support, incentives, and targeted information to build capacity to help leaders and educators improve.

Design Recommendations and Rationale

While the current goals for next generation accountability systems are largely upheld, the movement toward more student-centered, personalized learning environments suggests new considerations for the design and evaluation of accountability systems toward meeting these goals. Therefore, the Committee suggests that next generation accountability systems meet the stated goals through the following design recommendations:

Recommendation 1: Accountability systems should include a broad range of indicators that better capture the full construct of college, career, and civic readiness.

Explanation: In an accountability system that prioritizes college, career, and civic readiness, it is important to include indicators that a student is prepared for post-secondary transition or is ‘on-track’ to meet this expectation. There are numerous potential indicators for this category, particularly when one considers that readiness is a multi-faceted dimension that goes beyond academic performance and includes such characteristics as cognitive strategies, academic behaviors, and contextual skills (Conley, 2005). A framework for conceptualizing this construct developed by the ILN is presented below (CCSSO, 2013).

<p style="text-align: center;">Knowledge</p> <p style="text-align: center;"><i>Mastery of rigorous content and the facile application or transfer of what has been learned to complex and novel situations</i></p>	<p style="text-align: center;">Skills</p> <p style="text-align: center;"><i>The capacities and strategies that enable students to learn and engage in higher order thinking, meaningful interaction planning for the future</i></p>	<p style="text-align: center;">Dispositions</p> <p style="text-align: center;"><i>Socio-emotional skills or behaviors that associate with success in college, career and citizenship</i></p>
<ul style="list-style-type: none"> • Common Core State Standards • Career & Technical Education • Other Content Areas & Essential Literacies • Global Competence • Applied Knowledge 	<ul style="list-style-type: none"> • Critical thinking • Problem solving • Working collaboratively • Communicating effectively • Metacognition & self-awareness • Study skills & learning how to learn • Time/goal management • Creativity & innovation 	<ul style="list-style-type: none"> • Agency (Self-efficacy) • Initiative • Resilience • Adaptability • Leadership • Ethical behavior & civic responsibility • Social awareness & empathy • Self-control

Reproduced from: CCSSO (February, 2013). *Knowledge, Skills, and Dispositions: The Innovation Lab Network State Framework for College, Career, and Citizenship Readiness, and Implications for State Policy*. Available at: http://www.ccsso.org/Resources/Publications/ILN_CCR_Framework.html

Because it is often challenging to obtain reliable measures that are not easily corrupted (e.g. engagement measured via self-reports), some of all of these elements may not be suitable as factors that contribute to high-stakes outcomes. However, it remains important to promote and

track these indicators in the system. States may consider new assessment models for school and district accountability, such as matrix sampling, that provide information on system performance for a fuller range of CCR knowledge, skills, and dispositions. Such models could leverage the use of performance assessments or other measures of skills and dispositions that are deployed locally to support student learning, without tying such measures to high-stakes accountability at the individual student level, and without increasing the assessment burden.

Rationale: Systems can better incentivize and measure the goal of readiness by including a broader range of indicators that extend beyond K-12 and more authentically reflect the range of attributes and skills that are valued. Naturally, this will provide information to help teachers and leaders identify and address strengths and weaknesses. Moreover, the measures will serve as a clear signal to students and others (e.g. leaders, parents) about the nature of performance that is expected. By both signaling intent and providing a broader range of indicators against multiple dimensions of CCR, accountability systems can incentivize the development of strategies to support these outcomes.

Recommendation 2: Districts or schools should have flexibility to establish at least some priority outcomes for which they will be held accountable.

Explanation: Each state should establish rigorous statewide measures of CCR (such as through Common Core-aligned assessments), but should also provide latitude for district innovation to expand on those measures to include additional indicators of CCR skills or dispositions deemed important by the local community. While the state typically defines all the outcomes and measures in the accountability system, an alternate approach involves the state differentiating between outcomes and measures that must be standardized statewide versus those that the districts will have flexibility to define and implement. An example of this distinction is shown in Table 1.

Table 1. Example of Differentiated State and District Indicators.

Statewide Standardized Outcomes/ Indicators	District/ School Determined Indicators
<p><u>Outcome:</u> Achievement in Core Academic Content areas</p> <p><u>Sample Indicators:</u> Percent of students meeting target performance level on state ELA, math, and science summative assessments</p>	<p><u>Outcome:</u> Demonstration of problem solving, collaboration, and communication skills.</p> <p><u>Sample indicators:</u> District developed performance assessments, successful completion of capstone project/ activity, completion of selected courses of certifications</p>

The state’s role, then, is to define the common and standard outcomes, to determine the priorities or process for identifying a range of additional valued outcomes, and to provide criteria and guidance to support implementation. The state may even supply a ‘model system’ to help support

districts. But this model is not proscriptive and ultimately the district assumes the role of working with schools to select the priority outcomes and methods of implementation.

Moreover, states will determine the influence of these outcomes in the accountability system. Some indicators may be used for reporting, while others may contribute to the overall school outcomes/ classifications. It is expected that the school/district determined indicators will have a meaningful influence in the overall system and inform improvement and support strategies.

Rationale: Schools may justifiably privilege additional outcomes or rely on different methods to achieve these outcomes. Consider that alternative schools, charter schools, magnet schools etc. often have a distinct mission and focus area that should be addressed in accountability. This is not in conflict with a shared focus by all schools on performance and progress in core academic areas. Rather, this flexibility addresses additional valued outcomes that are difficult to standardize statewide.

Additionally, affording districts and schools the flexibility to develop and/or adopt indicators incentivizes districts and schools to innovate and develop promising approaches. For example, the very act of collaborating to develop and evaluate customized, rich performance measures increases the understanding of and emphasis on these outcomes in schools, which promotes the goals of achievement, equity and effectiveness. Finally, permitting some customization in accountability produces more authentic information to evaluate and support schools.

Recommendation 3: Accountability systems should permit flexible testing approaches at the point of student readiness.

Explanation: The move to more flexible testing should be expressly and clearly permitted to support transitions in teaching and learning to more personalized, competency-based pathways. Whereas current accountability systems primarily produce a single, summative outcome at the end of the year to inform school classifications, improved systems should promote collecting information throughout the academic year to inform teaching and learning on an ongoing basis to support continuous improvement.

This principle is particularly important for implementation of Competency Based Pathway (CBP) models. CBP models are based on the concept that student progress should be defined by mastery of a clear set of expectations or essential competencies. Such competencies represent the essential outcomes important for student to advance toward readiness for college and careers. Each essential competency should be mastered before the next set is introduced and students should be permitted to progress when readiness is demonstrated. The CBP approach calls for flexibility for student advancement that is not hindered by seat time requirements and/or limited opportunities for assessment. This design principle is not viewed as incompatible with requirements for annual determinations of performance (in fact, some states already take an integrated approach), which can be accomplished by aggregating at a specific point in time (e.g.

end of grade eight) a series of assessment components or modules administered flexibly at the point of readiness.

Moreover, this approach to assessment is not viewed as in conflict with state assessments currently in place or being developed to be aligned with the Common Core State Standards (CCSS), notably PARCC and Smarter Balanced. Rather, a CBP can be part of a comprehensive approach to assessment in which different information is collected over time and for different purposes.

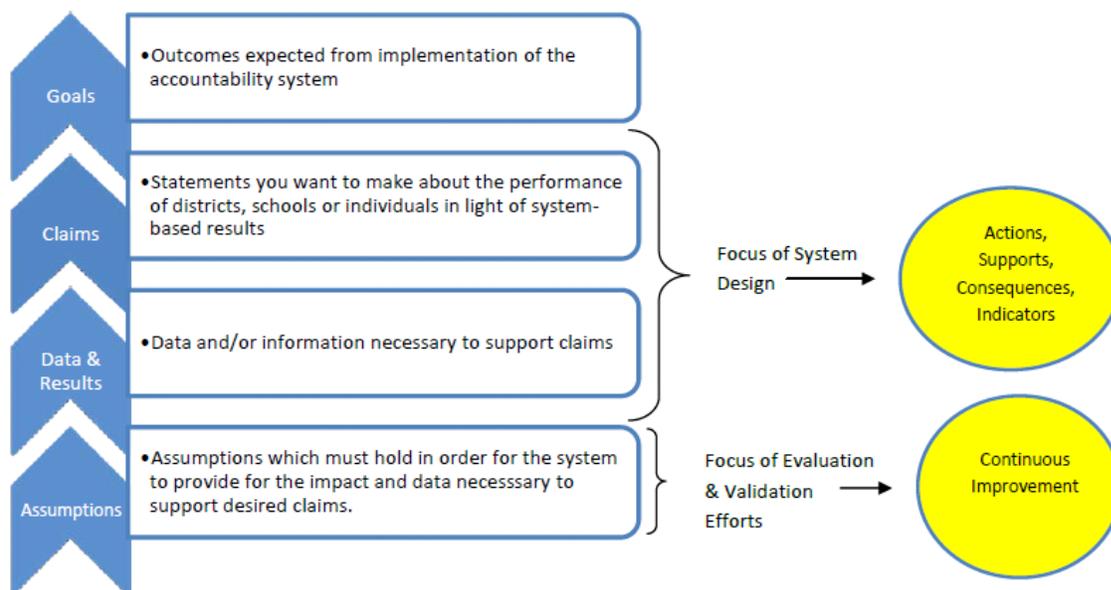
Rationale: The focus of this design principle is to promote practices that create more personalized instruction based on information collected through different approaches to assessment and build student agency in the learning process. Allowing flexible pacing permits students the time to get the focused support they need or aspire to higher levels of challenge. Additionally, more frequent and focused feedback improves the ability for teachers and students to concentrate efforts on the highest priority learning targets. This principle helps balance the attainment of academic achievement/growth and equity.

Recommendation 4: Accountability systems should be designed for continuous improvement, with a clear through-line between design features and improved student achievement.

Explanation: Accountability systems should not be designed to remain static, but should be designed with a clear theory of action (TOA) that explicates policymakers' assumptions about how the accountability system will bring about desired changes, and therefore provides means for evaluating the system for the purpose of continuous improvement. The theory of action should articulate how the goals of the accountability system and its underlying assumptions drive decisions around what data and information will be collected in order to assess progress toward those goals, as well as the delivery of supports and interventions. The theory of action should also address who will be responsible for the actions and how they are intended to be implemented.

A high level overview the components and connections that are should be included in a theory of action is depicted in Figure 1.

Figure 1. High Level Elements of a Theory of Action.



Rationale: A TOA provides a basis for evaluating the extent to which accountability design decisions support the attainment of stated goals. In particular, the TOA explicates how the actions, consequences and supports work together to build capacity and effectiveness. By including claims that are clear and potentially falsifiable, leaders can check that the system is functioning as intended and make improvements as necessary.

The TOA also helps ensure that multiple elements of the system are coherent. For example, it is important for states to have coherence among curriculum, instruction, assessment; and among student, school, district, and educator accountability. The system should provide incentives that are consistent and mutually supportive. For example, if an educator evaluation disincentivizes collaborative practice, but the school accountability system relies on such collaboration for success, the system lacks coherence. By explicating assumptions about how change is thought to occur, the TOA can help ensure the system functions coherently.

Conclusion

The central proposal reflected in the recommendations described in this document is that the top-down, standardized, state accountability systems that characterized NCLB should be transformed. Next generation accountability systems should focus on broader set of outcomes to better support post-secondary readiness and empower states, districts, and schools with the flexibility to construct and evaluate systems that will promote these outcomes. This vision document is meant to provide the initial information for continuous improvement and states leadership to ensure CCR outcomes for all students. Building on the leadership in these areas to date and with appropriate policy support and capacity, state leaders can and will continue to make improvements in developing innovative and effective education accountability systems.

**Performance Assessment
For Competency Education
(PACE)
District Guide for Planning & Implementation**

**New Hampshire Department of Education
September 2014**

**NH Department of Education
Performance Assessment for Competency Education
Guidance Document**

Table of Contents

1. Introduction	XX
2. Overview of the PACE Accountability System	XX
3. The Role of Competencies	XX
4. PACE Performance Assessment Components	XX
5. Personalization	
6. Educator Evaluation Connection	XX
7. Building Understanding and Support in the Community and Stakeholders	XX
8. Reporting to the Community	XX
9. District Peer Reviews	XX
Appendix A: District Planning Template	
Appendix B: District PACE Needs Assessment	
Appendix C: New Hampshire Work Study Practices	

**NH Department of Education
Performance Assessment for Competency Education
Guidance Document**

1. Introduction

The New Hampshire Department of Education’s Performance Assessment for Competency Education (PACE) accountability option for demonstrating student, school, and district progress enables participating school districts to attain two goals. First, it allows districts to use what we know about effective learning to support accountability determinations in the New Hampshire State Model Competencies and Work-Study Practices. Second, it shifts the locus of accountability from external to internal, from the state to the district. Rather than having the state evaluate the district using external measures and then tell the district how it is faring, the district will evaluate its own performance and then tell its own story within certain constraints, with accompanying evidence, to the state. These two goals represent a fundamental shift in how accountability is conducted. This first year of a two-year pilot represents a collaborative partnership to field test the PACE procedure and protocols. During the first year there will be districts piloting PACE (*implementing* districts) and those preparing to do so in the 2015-16 school year (*planning* districts). Upon completion of the pilot, the hope is that a sustainable PACE option will exist for any district who wishes to use. We are grateful to all partner districts for entering into this work with us.

The purpose of this guide is two-fold:

- 1) To define clearly the guidelines of what comprises successful participation in PACE.
- 2) To provide advice, support, and reference for districts making the shift to PACE.

The Guide is meant to explicate the guidelines for using performance assessment to support accountability reforms in New Hampshire. In some cases, the guide will provide specific information on technical requirements – such as the role of locally validated assessments in the system. In other cases, it will provide information on facets that are not required – such as communication with stakeholders – but which should increase the likelihood of successful implementation.

2. Overview of the PACE Accountability System

The PACE accountability option provides districts with an alternative route for demonstrating measurable progress in student outcomes relative to the New Hampshire state model competencies, Work-Studies Practices, and other important measures. It supports districts in emphasizing meaningful content, high quality instruction, and deep student engagement. The PACE accountability system will have multiple components, but performance assessment will be a central feature. In this first year, PACE districts will develop plans to report on ELA, math, and science, and the Work-Study practices. As New Hampshire develops further model competencies in social studies, the arts, and other content areas, the accountability system will likely increase in scope to include these content areas.

2a. The Purpose of PACE

In the PACE option, the New Hampshire Department of Education (NHDOE) has created a route for districts and schools to demonstrate progress that is not solely or primarily dependent upon state standardized tests. *The creation of the PACE accountability option reflects NHDOE's belief in placing the decisions of how accountability is determined in the hands of districts and schools, rather than the state mandating a uniform set of accountability measures.*

Within a context of reciprocal accountability, *districts and schools are responsible for determining and reporting on local accountability measures, while the state is responsible for providing assistance to districts in setting up strong accountability systems.* PACE is based on the belief that most students are better able to demonstrate their competency through multiple performance assessment measures in a variety of contexts. Thus, the accountability option was established to enable schools and districts to demonstrate student achievement and learning growth through means other than or in addition to standardized tests, with an emphasis on performance assessment.

Performance assessment must be a central component of each district's PACE system, for both summative and formative purposes. A performance assessment is defined as follows:

Performance assessments are multi-step, complex activities with clear criteria, expectations, and processes that enable student to interact with meaningful content and that measure the depth at which students learn content and apply complex skills to create or refine an original product and/or solution.

New Hampshire is committed to implementing a philosophically coherent system. If the State is encouraging districts to embrace student agency in determining learning goals, then it only makes sense for the State to embrace "district agency" in establishing its own accountability goals.

2b. PACE Design Criteria

The PACE option is designed to balance the strengths and prerogatives of the state and the local district. School districts, with extensive support from the State, will design the accountability

system by which they will establish and monitor achievement and growth goals, while the State will establish criteria for these systems. These criteria reflect broad principles around equity and opportunity. Districts will be required:

- To commit the resources necessary to ensure the plan’s success;
- Exhibit leadership and educator capacity to design, implement, support, and sustain the system; and
- Demonstrate a clear and visible commitment towards improving the achievement of students who have historically been underserved.

Because each PACE district comes to this work with different levels of experience and areas of focus, the NH DOE recognizes that PACE will “look different” depending on the district. Still, the expectation for all *implementing* districts (as opposed to planning districts) is that by June of 2015 districts will present evidence of a PACE system sufficiently developed to implement fully in the 2015-2016 school year that addresses the technical quality criteria in section 4f below.

In 2014-2015, PACE *planning* districts will be expected to engage in beginning conversations about building capacity to implement a PACE accountability system; this may include the initial development of PACE performance tasks, professional development to build literacy assessment with district faculty, and mapping out gateway grades and tasks to assess demonstration of achievement of competencies. During 2015-2016, PACE *planning* districts will be expected to either (1) develop, pilot, and refine for tasks in English language arts, mathematics, and science for submission to the state performance task bank, while planning for full implementation in 2016-2017, OR (2) fully implement a PACE school and district accountability system.

The expectation is that eventually, each PACE district’s system of assessments will meet a set of criteria mutually agreed to among NHDOE and PACE districts, while taking into account USED requirements. Following is a draft set of criteria that will be used as a starting point:

A. Coherent Set of Multiple Measures

- Provides evidence of a coordinated district plan to assess and report on student performance
- Contains measures within, at least, the following domain:
 - Academic (e.g., performance tasks, Work-Study Practices, use of SB in one grade in each grade span)
 Additional recommended domains include the following:
 - College and career readiness (e.g., Advanced Placement, graduation, dropout)
 - Teacher effectiveness (e.g., educator evaluation)
 - School culture/environment (e.g., parent, student, teacher surveys)
 - Student engagement in learning (e.g., student surveys)
- Yearly measureable targets are established that are coherent and ambitious

B. System of Performance Tasks to Assess Student Learning and School Progress

- Selection and use of a combination of validated state and local competency-based performance assessment tasks in academic disciplines with state competencies

(currently English language arts, mathematics, science) which provide rich data on how students are progressing at every grade span

- For individual assessment tasks: alignment, fairness, and consistency and accuracy
- For the system of assessments: alignment, fairness, comparability, and consistency and accuracy
- Process of local task design and scoring that ensures task quality and scoring calibration
- Defines what is acceptable and unacceptable performance for each performance measure
- Local policies that support the state performance assessment system (e.g., performance assessment graduation requirements)

C. Opportunity for All Students

- Must demonstrate the opportunity for an adequate education for all students
- Instructional system is designed to support student learning of competencies
- Academic supports are in place to assist struggling students to meet competencies

2c. Evaluation of PACE

NHDOE is committed to supporting pilot PACE districts, and believes that accountability is reciprocal between the State and the Districts. Aside from ongoing policy work, the NHDOE will provide development to district teachers and leadership, facilitating cross-district task work, and tying in the work of the New Hampshire Performance Assessment network. The goal of this work is to provide assistance in developing and implementing the local accountability system. Additionally, PACE districts will be convened quarterly throughout the school year to share practices, problem-solve dilemmas, and be introduced to additional accountability tools and supports.

2c. PACE District Support:

NHDOE is committed to supporting pilot PACE districts, and believes that accountability is reciprocal. Aside from ongoing policy work, the NHDOE will provide development to district teachers and leadership, facilitating cross-district task work, and tying in the work of the New Hampshire Performance Assessment network. The goal of this is to provide assistance in developing and implementing the local accountability system. Additionally, PACE districts will be convened quarterly throughout the school year to share practices, problem-solve dilemmas, and be introduced to additional accountability tools and supports.

2d. District PACE Self-Assessment:

Districts engaging in the PACE option are expected to complete the District PACE Self-Assessment, and use the results from the assessment to guide discussions and planning in designing the local PACE accountability system (see Appendix). The purpose of the self-

assessment is to provide districts with a clear view of where their energy and resources should be directed in terms of meeting PACE requirements and setting priorities.

2e. Planning Template:

At the beginning of the 2014-2015 school year, districts are required to submit an action plan, developed with guidance from their PACE coach. A planning template is provided (see Appendix), but if the district has a preferred action planning model they are free to use it. The plan must include a series of goals – determined after reviewing the results of the self-assessment – and for each goal a set of action steps that lead to the meeting of that goal. Action steps should have clear time markers, and must designate a person who will be responsible for shepherding that action step to completion. In the planning template provided in the appendix, a column is provided to allow the district to consider what help might be needed from the NH DOE.

DRAFT

3. The Role of Competencies

In the PACE accountability plan, competencies are at the forefront of the design process of high quality performance tasks. To date, the New Hampshire Board of Education has approved three sets of state model competencies that will be used in the design of performance tasks as part of PACE – English language arts, mathematics, and science.

The NH Board-approved model competencies in English Language Arts, Math, and Science have all been through a vetting process. The competency validation process assures that the competency statement meets the ‘Strong’ criteria for enduring nature, cognitive demand, alignment to standards, and accessibility. This quality assurance process is meant only as a quality assurance tool that the competency statement is strong and of high quality. Strong competency statements facilitate the development of performance tasks.

These competencies are aligned to the Common Core Standards in English Language Arts and Mathematics. Because of this alignment, performance tasks designed at the local level should be mapped to these competencies. Rubrics for these tasks can also be tuned for proficiency by cross-referencing the grade-level expectations for performance embedded within the Common Core State Standards documents.

As performance tasks are designed by cohorts of teachers, designers should be using the competencies as a starting point for design, then determine how students will be assessed relative to the competencies.

In the event that designers are writing tasks using locally developed competencies instead of the state-approved competencies, it is important that each competency statement is validated using the Competency Quality Assurance Rubric on the NHDOE website.

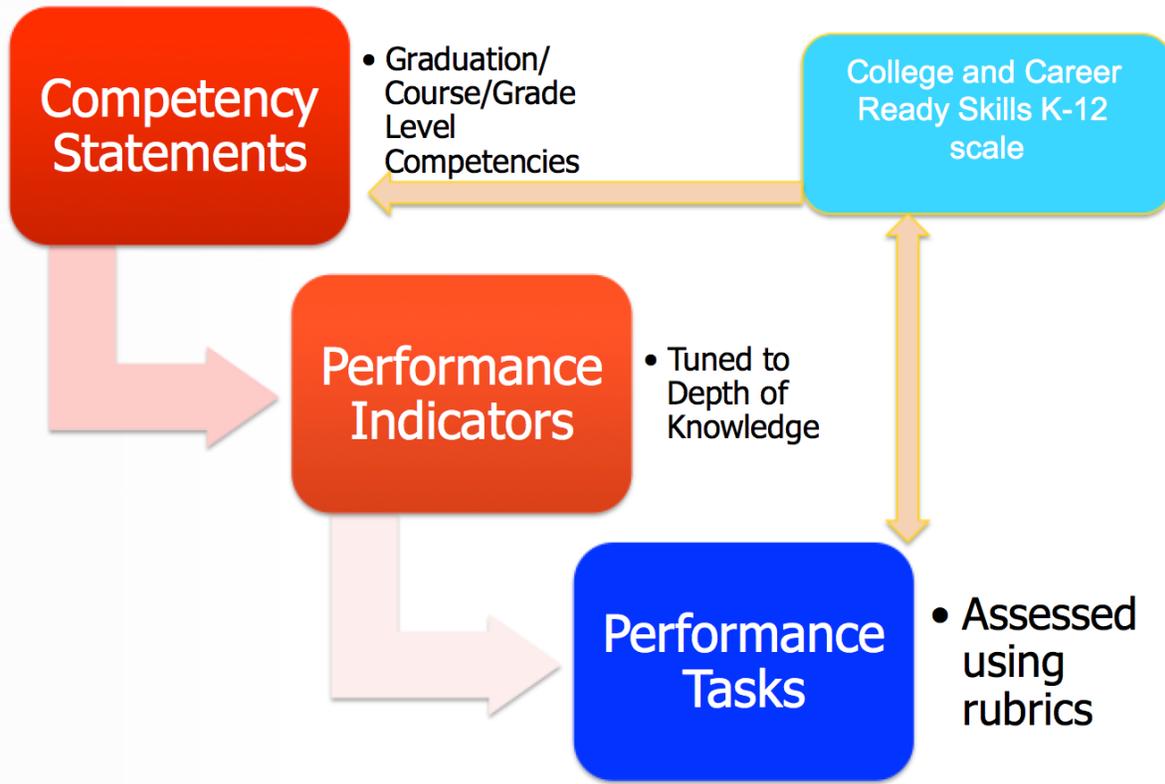
Any performance task for PACE should have complete validated competency statement(s). Each task, when unpacked, should contain performance indicators that can be mapped back to the demand of the competency.

Figure 1 outlines the process for this framework. Competency statements that have been designed and validated as part of a graduation course, or grade-level competencies are unpacked into a series of performance indicators of “I can...” statements. Each of these statements should be categorized by Webb’s Depth of Knowledge (Hess’ Cognitive Rigor Matrices can be a useful tool in these conversations). Each performance task should have embedded within it performance indicator (s) that meet the criteria for Level 3 or 4 on Webb’s Depth of Knowledge. This guarantees that the assessment performance represents “Proficient” or “Competent” for the identified Competency.

Also note in Figure 1 that the College and Career Ready Skills (Work-Study Practices in NH) should be embedded in the performance task. There should be appropriate rubrics for these Work-Study Practices used as assessment criteria in the performance task. The Work-Study Practices identified for NH include: collaboration, creativity, self-direction, and communication.

The draft Work-Study Practices rubrics and materials developed by the statewide work group are included in the Appendix.

(Figure 1. Competency Design Framework)



As an example of the process consider the following:

Competency statement:

Students will demonstrate the ability to investigate and analyze a natural or human designed system in terms of its boundaries, inputs, outputs, interactions and behaviors and use this information to develop a system model that can be used to understand and empirically evaluate the accuracy of models in terms of representing the underlying system.

Performance indicators:

Performance Indicator Statement	Depth of Knowledge
I know and can explain the function of each structure in an animal cell	DoK 1 or 2
I can describe cell receptor site activation	DoK 1 or 2
I can analyze and synthesize research information on the effect of insulin at receptor sites and its biochemical effect within a cell	DoK 3
I can design and build a 3D model of receptor site activation that replicates intracellular changes in response to receptor site activation. I can explain the decisions that led to my design and evaluate their efficacy.	DoK4
I can use the engineering design practices to design, build, and test my 3D model	DoK3

*Table 1***Performance Task:**

Following your research of the dynamic relationship between receptor sites on a cell membrane that results intracellular changes, your group will design and build a 3D model or receptor site physiology that depicts how a receptor site on a cell membrane causes a change within a cell. Create presentation materials for your model which may be created using materials or using 3D design software. (Rubrics for research, reflection, product, and presentation were created, validated, and tested for reliability.)

4. PACE Performance Assessment Components

Performance assessments used in a PACE system must be tightly aligned to the New Hampshire State Model Competencies and the Work-Studies Practices.

Every PACE district must have a rating scale for PACE performance assessments. Each district may have their own wording, but the levels must be defined and aligned in expectations. Each district will also need to develop common descriptions of proficiency across disciplines for each grade level that are aligned with state grade-level descriptions of proficiency (the latter to be developed during the 2014-2015 school year).

4a. Selection of Tasks

Districts are permitted control over the number and distribution of the performance assessments within their PACE accountability system, within certain parameters.

1. Performance assessments must provide evidence of, and report on, student demonstration of achievement of competency in ELA, Math, and Science (Social Studies and the Arts will be added as New Hampshire State Model Competencies in those areas are developed).
2. Performance assessments must provide evidence of achievement of the Work-Studies Habits.
3. Students must be allowed multiple opportunities to demonstrate evidence of achieving a competency over the course of a year.
4. Districts must use a mixture of locally-designed performance assessments and assessments drawn from validated state/multi-state task banks.
5. Districts must use AT LEAST ONE common performance assessment drawn from a validated state/multi-state task bank PER CONTENT AREA, PER YEAR in grades 3-8 and at least one high school (9-12) grade. In this context, a “common assessment” is one given to all students of a grade level or course in a district.

The vision of the PACE system is not for districts to assess for each competency with singular assessments that “prove” demonstration of achievement of a competency. Rather it is to allow for the collection of valid, sufficient, and credible evidence that, over time, allows a student to demonstrate achievement of a set of competencies.

The multiple assessments afforded students should allow for demonstration in multiple modes of product (e.g., written work, oral presentation, etc.), at a varying level of complexity. They should also allow for a variety of levels of student choice, with some performance assessments perhaps being more structured while others provide more latitude for the student determining how they will choose to demonstrate attainment of the competency.

4b. Task Administration

While there are few hard and fast rules regarding task administration of PACE performance assessments, the PACE vision is that tasks will provide sufficient and credible evidence for students to demonstrate attainment of competencies in a meaningful way over time. Tasks may

allow for all work to be done in class, or a mixture of in-class and out-of-class work, with, for example, research being done off-site but check-ins, feedback, and instruction being provided in class. PACE allows for a range from standard PAs to complex projects and student designed assessments (see Chapter 5, Personalization).

Performance Assessments should be embedded into the curriculum, and scaffolding should be provided as necessary to all students in ways that do not compromise the intended measurement target. For example, a student with low literacy levels taking a task that measures a science competency may have materials read to them, but the same student taking a task that measures a reading competency may not, if it is measuring a fluency aspect of reading.

Students who do not demonstrate attainment of a competency on a task should – after additional instruction – be provided with another opportunity to attempt the task, or another opportunity to demonstrate attainment of that competency in a different task.

4c. Local Task Design and Quality Assurance Process

Whether a task is valid or not is determined by whether the task measures the competency it purports to measure, and whether it does so in a way that is fair and accessible to all students. In addition to use of state common tasks, it will be necessary for districts to design and use locally developed performance tasks as part of their PACE accountability system. In order to assure the quality of these tasks the district must develop or adopt a system-wide design and quality assurance, using the NH QPA Task Quality Assurance Protocol. The criteria must include

1. Alignment to New Hampshire Competencies and Work-Study Practices
2. Level of Cognitive Rigor
3. Fairness and Lack of Bias
4. Rubric Design

The quality assurance process must include a sequence for each task or design that includes initial quality assurance, field testing, analyzing student work, revision, further field testing, and further analysis of student work. Task design and quality assurance is an iterative process and must be reflected in the district's process. A number of tools can be used to assure quality throughout the development and implementation of a task, including the Task Quality Assurance Checklist, a calibration protocol, a task design template, and a number of other QPA tools that have been refined for the NH DOE specifically for PACE use.

4d. Calibration and Inter-Rater Reliability

Along with the validity of its tasks, a district must develop a system-wide process for assuring that tasks are administered and scored consistently. Inter-rater reliability is a measure of the consistency with which two or more raters consistently score student work. Within PACE, inter-rater reliability needs to occur within districts and across districts.

A process assuring inter-rater reliability will require that teachers calibrate with each other before scoring common tasks, and then analyze student work and scores together afterwards.

Districts should use the NH QPA Calibration Protocol for this process. This can be especially valuable when teachers from multiple districts gather to calibrate and analyze each other's work. Opportunities for such cross district calibration and analysis will be provided during the 2014-2015 school year. If districts want to initiate their own partnerships and collaborations, they are encouraged to do so.

4e. Technical Criteria

The degree to which the performance assessments are useful for determining whether students have attained specific competencies is contingent upon the technical quality of each assessment. However, since it is doubtful that individual competencies can be assessed with a single assessment, the technical quality of the system of assessments is even more important.

While there are a number of technical criteria that can be used to evaluate the performance assessments used as part of the PACE system, we have decided to focus first on the highest leverage criteria. *The criteria outlined below, if met, would ensure that students are being provided with a high quality assessment experience.* There are three main criteria for **individual assessments** and four criteria for the **system of assessments** (discussed below), with obvious overlap among the two sets of criteria. The descriptions that follow represent expectations for a fully developed PACE system. It is assumed that PACE districts will progress to this level of technical quality over a period of years. The staging of this progression will be discussed in the Peer Review section below.

The following three criteria for evaluating the technical quality of **individual assessments** are all intended to support evaluations of the validity and generalizability of the scores from the assessments:

- ✓ Alignment
- ✓ Fairness
- ✓ Consistency and Accuracy

Alignment is an evaluation of the degree to which the assessment reflects and is designed to measure student learning of the specific standards and/or competencies at the level of cognitive rigor indicated in the competency. **Fairness** is the degree to which the assessment is comparably valid for various groups of test takers. **Consistency** is the degree to which a student will achieve the same results if we could administer the same hypothetical assessment an infinite number of times to students, have it scored by an infinite number of raters, and if we could administer all possible tasks purportedly designed to measure the same aspect of the particular competency. **Accuracy** is the degree to which the assessment reflects the true achievement of students or in terms of scoring, the degree to which scorers are able to score the same as expert scorers.

These brief descriptions of the criteria are not enough to guide evaluations. Even more detailed explanations are not necessarily helpful to district leaders as they work to ensure that their assessments meet these critical criteria. Therefore, in addition to a more elaborate description, we offer suggestions for the types of evidence that could be used to document the degree to which a district's assessments are meeting the specific criteria. The table below outlines the criteria,

specific descriptions of each criterion, and examples of evidence that can be used in support of the criteria.

Criterion	Specifications	Example of Evidence
Alignment	<ul style="list-style-type: none"> a. The content and skills elicited by the task represents key content and skills described by the competencies. b. The key content and skills elicited by the task are assessed by the scoring rubric. c. What is elicited from students on the task requires the depth of knowledge expected of students as stated in the relevant competency. d. The depth of knowledge expected of students is assessed by the scoring rubric. 	<ul style="list-style-type: none"> a. A description of the degree to which the content, skills and depth of knowledge expected by the relevant competency are elicited by the task. This should include exemplar student responses. b. A description of how the scoring rubric assesses the content, skills and depth of knowledge expected by the relevant competency. This should include exemplar student responses. c. A description of the process used to evaluate that each task is matched to a specific competency and that each scoring rubric assesses what is expected by the competency. This should include a rationale for the process used.
Fairness	<ul style="list-style-type: none"> a. The task does not contain excessive language demands including unusual words, wordiness, and irrelevant information. b. Material is accessible to students from identifiable cultural, gender, linguistic, and other groups. 	<ul style="list-style-type: none"> a. A description of the degree to which the vocabulary and context presented by the task is free from cultural or other unintended bias. b. A description of how the task uses appropriate levels of academic language for the grade and content area. c. A description of the processes and people used to review tasks for issues of fairness and accessibility and the rationale for why these processes were appropriate.

<p>Consistency and Accuracy</p>	<p>a. The scoring rubric designed for the task will yield consistent scores given that raters have been trained appropriately.</p> <p>b. For multidimensional tasks, scores on each dimension should be within one point of other raters at least 90% of the time.</p> <p>c. For multidimensional tasks, overall scores should be within two points of other raters at least 85% of the time.</p> <p>d. For multidimensional tasks, scores on each dimension should be within one point of an expert rater at least 90% of the time.</p> <p>e. For multidimensional tasks, overall scores should be within two points of an expert rater at least 85% of the time.</p>	<p>a. A description of the scoring procedures for individual tasks.</p> <p>b. Data demonstrating that inter-rater agreement for individual tasks meets acceptable levels of agreement.</p> <p>c. Data demonstrating that agreement between raters and experts for individual tasks meets acceptable levels of agreement.</p> <p>d. A description of the plan for monitoring inter-rater agreement and agreement between raters and experts for individual tasks.</p>
--	--	--

Table 3

4f. Technical quality of the assessment system

As noted above, it is important to have high quality individual tasks and assessments, but doing so does not guarantee that the district will have a high quality assessment system. A thoughtful and coherent system of assessments is necessary for evaluating student achievement of the required competencies. In other words, a district might have eight very high quality assessments that it has employed to evaluate high school mathematics competencies, but if all eight assessments are focused on only a narrow slice of the competencies, we would not have a valid assessment system. So while many of the same criteria are used for both individual and sets of tasks, the focus of the criteria for the system of tasks is on the full system. For example, the “alignment” criterion for individual tasks/assessments is intended to evaluate the degree to which an assessment appropriately represents the specific learning target represented by a certain aspect of the competency. However, the same criterion at the system level focuses on the degree to

which the set of assessments represents the full set of competencies for the specific grade level or span.

The following four criteria for evaluating the technical quality of **systems of assessments** are all intended to support evaluations of the degree to which determinations of competencies determinations can be validly made from the assessments:

- ✓ Alignment
- ✓ Fairness
- ✓ Comparability
- ✓ Consistency and Accuracy

Comparability is the degree to which the results of assessments intended to measure the same learning targets produce the same or similar results. In a psychometric view, true comparability is when the scores of one assessment are completely interchangeable from the scores on another assessment. In other words, a score of 240 on assessment A would lead to the same score inference as a score of 240 on assessment B. This is too strict of a standard for most locally-developed assessments, but that does not mean we cannot create assessments that lead to similar score inferences. For the PACE system, comparability at the system level is especially important. For example, we want to ensure that students within and across schools are being held to similar expectations within and across years.

Criterion	Specifications	Expected Evidence
Alignment	<p>a. The set of tasks fully represents the content and skills of the competencies.</p> <p>b. The set of tasks fully represents the depth of knowledge expected by the competencies.</p>	<p>a. A test blueprint that includes a label clearly identifying the specific tasks used, the competencies assessed with each task and the depth of knowledge assessed with each task.</p> <p>b. A narrative describing the processes used to evaluate alignment among the competencies and the set of tasks. This should include a description of the process used to ensure that each competency has been adequately assessed by the set of tasks.</p> <p>c. A brief description of how alignment was considered as part designing the set of tasks.</p>
Fairness	<p>a. Appropriate accommodations are used</p>	<p>a. A description of policies and procedures used to ensure the system</p>

	<p>so that educators are able to make valid inferences about the achievement of students with disabilities and English language learners.</p> <p>b. The set of tasks incorporates a variety of task formats so that students who tend to perform better or worse on specific formats are not treated unfairly.</p> <p>c. Assessment conditions are appropriate for all students.</p> <p>d. Tasks that reveal differences in mean scores among subgroups are compared to the results on other tasks and external measures (e.g., SBAC and NECAP).</p>	<p>does not unfairly disadvantage any group of students. This should include test administration protocols.</p> <p>b. Policies and procedures for ensuring fair participation of all students in the system regardless of disability or English language proficiency.</p> <p>c. Evidence that accommodations and alternate assessments are used as intended. This include documentation that the accommodations noted in a student’s IEP/504 plan were appropriately provided.</p> <p>d. A description of how the system provides students with multiple opportunities to demonstrate their knowledge and skills. This description is not of each individual task, but rather of how within the set of tasks different formats and strategies are provided.</p> <p>e. Statistical analyses of mean score differences among subgroups and relevant comparisons with other tasks and external measures (e.g., SBAC and NECAP). This includes an investigation if comparisons yield substantial differences.</p>
<p>Comparability</p>	<p>a. If tasks require the same knowledge, skills and abilities and if the scores they yield are equally accurate for students at any given level of those</p>	<p>a. A description of the process used to ensure the comparability of assessments across the schools and classrooms within a district.</p> <p>b. A description of the process used to</p>

	<p>knowledge skills and abilities, then the scores they yield should be comparable.</p> <p>b. Tasks are related to other measures thought to measure the same or very similar learning targets.</p> <p>c. The set of tasks produces comparable results across schools and classrooms within the same school district within a given year and across years.</p>	<p>revise and replace assessment tasks and how it is ensured that the new tasks represent the same content, focus, and depth of knowledge as the tasks that are being replaced.</p> <p>c. Documentation that the system requirements are comparable for students within each district both within a given class cohort and across years.</p>
Consistency and Accuracy	<p>a. The overall generalizability of the set of tasks for a given competency is sufficient for the intended purposes.</p> <p>b. The system includes carefully constructed scoring rubrics, appropriate training for raters, and a process to monitoring the scoring of student products.</p>	<p>a. A description of scoring procedures for the set of tasks. This should include procedures used for training raters and monitoring the scoring of student products.</p> <p>b. A description of how the overall generalizability of the set of tasks for a given competency is sufficient for the intended purposes.</p>

Table 4

4g. The Role of Smarter Balanced Assessments in the PACE system

All PACE districts will continue to participate in the Smarter Balanced Assessment system by having students complete the end-of-year summative assessments¹. Further, Smarter Balanced will be offering a variety of interim assessment options for states and districts. The specific nature of the full suite of Smarter Balanced is still being determined, but the NH DOE will be working with participating districts to identify appropriate uses of the Smarter Balanced interim

¹ The specific nature of the required participation (e.g., every grade, 3-8 and 11 or some selection of grades) is currently being negotiated between the NH DOE and the U.S. Department of Education.

system as more information is available². Ideally, Smarter Balanced will produce interim tasks that can be used to support determinations of student competency, but these will have to be mapped to specific competencies to determine the alignment with the NH State Model Competencies. These will likely be high quality assessments so districts are encouraged to find ways to incorporate these tasks and interim assessments.

Several questions have been raised about the role of Smarter Balanced summative assessments in competency education systems. While the end-of-year test is at too large of a grain size to make determinations about attainment of competencies for individual students, they provide useful information for auditing/monitoring at the system level. For example, while the results of Smarter Balanced summative assessments will not be perfectly correlated with local competency-based assessments, there should still a positive relationship at the aggregate level when comparing Smarter Balanced summative results with local assessment results. To the extent that local assessment results offer a discrepant picture of student achievement than Smarter Balanced assessment results, the district will be expected to evaluate and try to explain such discrepancies.

4h. Aggregating evidence making competency determinations

One of the biggest challenges in any accountability system—and competency-based education is a student-level accountability system—is determining approaches for aggregating and weighing evidence and then using the aggregated evidence to establish “mastery” or other identified performance levels tied to an inference of “competence.” We base the following discussion on the assumption that competencies represent big enough ideas that they cannot be validly assessed with just a single assessment. Therefore, it is likely that the results from the multiple assessments will need to be aggregated in a defensible way in order to determine if students have achieved the required performance on the competency.

There are many approaches for combining multiple indicators to yield a single outcome: *compensatory*, *conjunctive*, *disjunctive*, and *profile* methods.

- **Compensatory** means that higher performance in one measure may offset or compensate for lower performance on another measure.
- **Conjunctive** means that acceptable performance must be achieved for every measure (e.g., AYP).
- **Disjunctive** means that performance must be acceptable on at least one measure.
- A **profile** refers to a defined pattern of performance that is judged against specific performance level descriptions. A profile approach is often operationalized using a matrix to combine indicators for making judgments.

² Further, Paul Leather (NH DOE) and Scott Marion (Center for Assessment) are serving as two members of an expert panel examining the potential uses of Smarter Balanced assessments for competency education determination. The report, which will be released in the Fall of 2014, will also include recommendations for additional assessment options that Smarter Balanced should consider to better support competency determinations.

Given the challenges involved in characterizing the complexities associated with making determinations of competency, districts must employ a thoughtful approach for combining the multiple sources of data in order to arrive at the most valid inferences of competence.

It is beyond the scope of this document to offer a detailed discussion of the complexities of combining multiple indicators, but it is important to point out the risks associated with simple averaging techniques for determining student competence. An average (which is the most common compensatory approach) allows considerable variability in performance across multiple assessments to go unnoticed in the overall rating. Districts will need to develop an explicit plan that connects their specific assessments to the approaches for aggregating the assessment results for each competency.

Concurrent with the plan for aggregating data, districts must consider how they will set standards (i.e., cutscores) for making determinations of student competency. Standard setting is the process of converting written descriptions of performance into cutscores on an exam or other collection of quantitative data. It is a deliberate and systematic process designed to develop shared meaning among those selected to participate in the standard setting activity (the panelists) so they can establish cutscores based on a common understanding of performance. Technical assistance will be provided to participating districts to help them wrestle with such issues.

5. Personalization in Performance Assessment

Throughout this Guide, information on how to make performance assessment as comparable as possible – from teacher to teacher, building to building, district to district -- has been stressed, as well as the importance of fidelity to the model and common use of the statewide competencies. It is equally as important to design student voice and choice into the *overall* assessment system of any school and district. According to Andy Calkins, Deputy Director of Next Generation Learning Challenges, personalization is defined as

Students' learning experiences - what they learn, and how, when, and where they learn it - are tailored to their individual needs, skills, and interests, and enable them to take ownership of their learning. Although where, how, and when they learn might vary according to their needs, students also develop deep connections to each other, their teachers and other adults. (www.nextgenerationlearning.com)

In other words, the rigorous and deep learning goals and high expectations are, ultimately, the same for each student, while the learning environment and pathway responds to and adapts to each student's learning needs.

We have examples of this in New Hampshire's implementation of Extended Learning Opportunities. These are learning experiences, for credit, in which a student can engage outside the walls of the traditional classroom, connecting with content experts in the wider community, with mentors who are successful in the area that student is exploring and/or with peers who are engaged in the same deep exploration. Students' learning is developed around both content and work-study practice competencies that form the learning trajectory, but how far the student can soar is up to the student. Extended Learning Experiences are assessed through performance assessment and can represent what Joe DiMartino, Executive Director of the Center for Secondary School Redesign, refers to as "common assessment of uncommon learning."

These experiences have shown that personalized learning is a viable and reliable option for the classroom engaged in quality performance assessment. An assignment that was formerly a "paper" with detailed guidelines for students to follow can now become a "project with a written component" in which each student has wide choice of topic, format, style of presentation of material, and how extra information outside of class or textbook learning is garnered (online or community-based research, interviews, surveys, etc.) – all while working toward the same core competencies that are required for success in that subject. When learning is based on common competencies, it remains rigorous and focused even when individual students take various pathways toward those goals.

Why consider the quality of personalization in learning when designing your overall performance assessment system?

- Personalization is good for students who are dis-engaged.*

Over time, as schools in New Hampshire implemented personalized learning, a remarkable phenomenon was noted. Students who were previously disengaged - at risk for leaving school for personal reasons, at risk for failing out of school or out of more than one class, and students who had dropped out of school - but who agreed to return for a personalized learning experience were reaping great benefits from these types of learning experiences. In many cases, the students' attitude towards and attachment to school showed a complete turn-around; students re-engaged in a full school load and were successful in classes – even those that were not personalized. They and their teachers attributed this transformation in learning to the personal self-efficacy students experienced and to the attachments to adults in the school that they had forged through the personalized learning experiences.
- Personalization is good for students who fly ahead of their age peers.*

All schools and classrooms need options for students who forge ahead of the planned curriculum or who need deeper, wider, and faster learning experiences. When learning is based on competencies that are common in terms of the ultimate goals of learning, there is no limit to what students, with their educators' guidance, can design in order to keep accelerated students excited and engaged. Because personalization intrinsically includes a high degree of student ownership of learning, teachers are able to give these students the freedom to pursue their own learning at different rates and using different methods.
- Personalization is good for students with special needs.*

Personalization in learning is about effective teaching. It requires flexibility to follow the student's needs, interests, and goals. It requires differentiation in both pace of and approach to learning – moving beyond a one-size-fits-all type of education. Students with special needs shine in this type of learning environment, so competency-based learning that is personalized allows students of all learning styles to learn together, even as each accesses the specific resources or extra help that they need for their pace and their style.
- Personalization is good for all students, good for educators, and good for schools.*

One of the most promising aspects of personalization based on competency attainment is that it allows students to demonstrate their mastery of knowledge, concepts, and skills in ways that are unmistakable. It frees students to move on with complex learning when they are ready and frees teachers to devote their time and attention to those students who need extra help. M.B. Horn in a 2014 Education Next article said, "...when we move to a competency-based learning system concerned with rigor—in which students move on to new concepts only upon mastery (and there exists the notion of a minimum pace so students who are falling behind get more attention and gaps don't grow too big)—that students who would typically be left behind and see their gaps grow bigger and bigger, instead experience a sea change when misconceptions are corrected, they master foundational knowledge and skills, and they can then accelerate much faster than anyone would have expected."

<http://educationnext.org/stop-false-generalizations-personalized-learning/>

The nature of performance assessment, with the picture of the whole student that emerges over time, goes hand-in-hand with personalized learning both in the classroom and in extended learning experiences. Performance assessment as described within this guide includes both formative and summative assessments. While for accountability purposes, some learning tasks used for summative assessment will be common and may include a higher degree of prescription, the learning experiences and the attendant formative assessments that lead to these common summative assessments can be personalized. Schools that are attentive to student engagement and ownership of learning see that teaching moves to a more personalized approach. Such schools realize the true potential of competency education.

DRAFT

6. Educator Evaluation Connection

The role of performance assessments in Student Learning Objectives

Competency-based education is not the only accountability system that districts have to design and implement. All New Hampshire school districts are in the process of designing and implementing local educator and leader evaluation systems in addition to responding to school accountability requirements from the state. Many wonder if having to set up competency-based education and educator evaluation systems requires “double work.” There is no question that it requires more effort to design and implement two systems compared to one, but the systems should be designed to maximize coherence and efficiency. One significant area of overlap between the two systems is with the use of Student Learning Objectives (SLO) for documenting teachers’ contributions to student learning.

SLOs include a learning goal, targets for student and educator performance, and assessments designed to measure student learning of the learning goal³. The learning goal should reflect a big idea of the discipline, much the way that a competency statement reflects a big idea of the discipline. The assessments of the student competency, assuming they are appropriately rich and high quality assessments, could (and should) be the same assessments used to evaluate the learning goal of the SLO. SLOs require defining expected levels of student performance on the learning in terms of student targets and evaluated through the use of high quality assessments. These could and likely should be the same targets used to define acceptable levels of performance for students against the competency. Finally, the SLO requires an aggregate target (e.g., how many students meet their targets) for judging educator performance. This target is closely linked to the student targets, but other than that, does not have an analogous component in the competency system. Figure 2, below, provides a graphical representation of this relationship.

³ For more information on New Hampshire’s state model educator support and evaluation system including the use of SLOs, see: <http://www.education.nh.gov/teaching/documents/phase2report.pdf> or the Center for Assessment SLO Toolkit: <http://www.nciea.org/slo-toolkit/>.

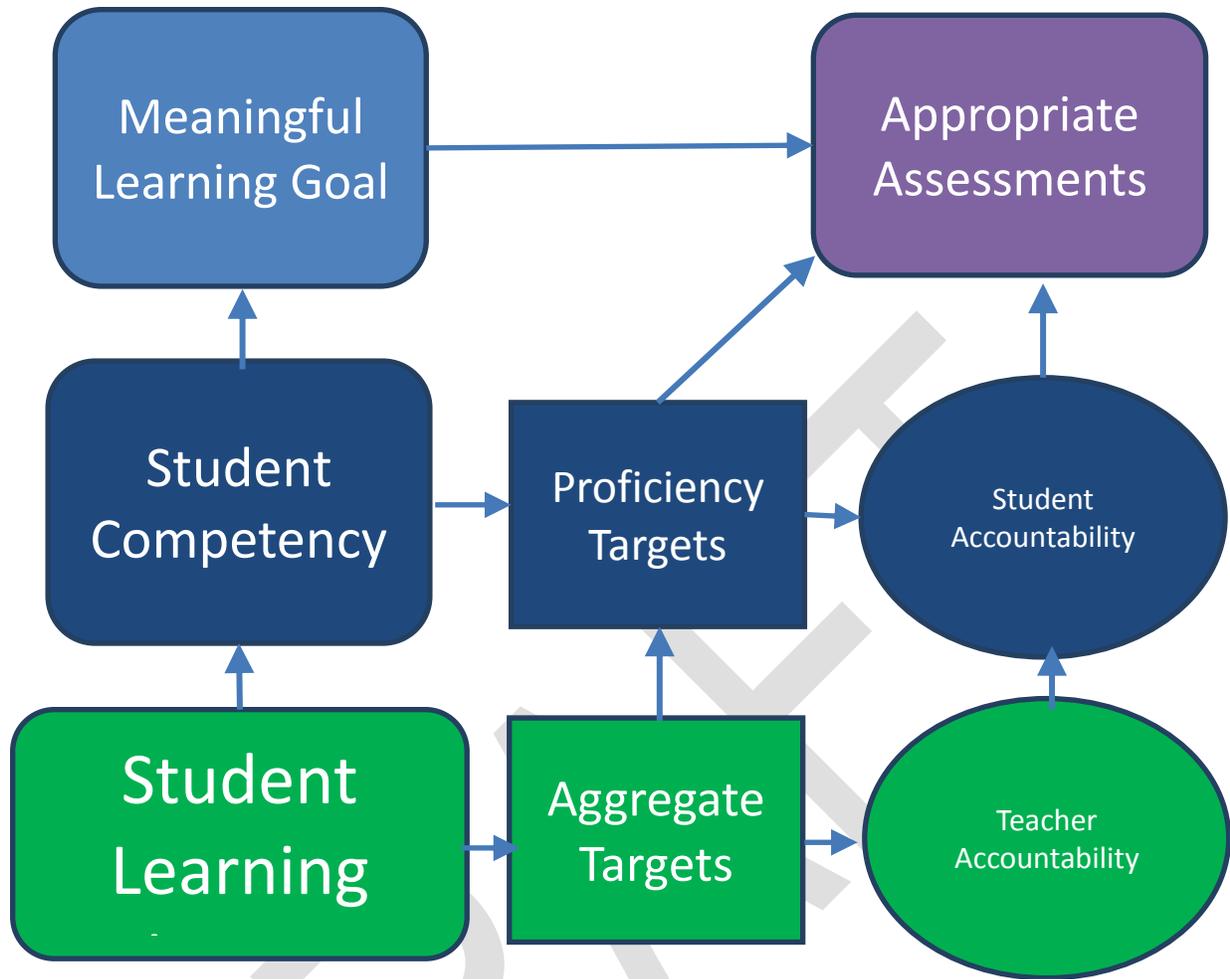


Figure 2. The coherence between student competency determinations and student learning objectives.

7. Building Understanding and Support in the Community and Among Stakeholders

PACE districts are required to develop a communications plan that encompasses all stakeholders, including students, teachers, administrators, parents, community members and school board. The specifics of the communications plan will be determined by the district and must engage the stakeholders in a conversation that educates them and fosters support for the district's PACE accountability plan.

A part of this plan will be formal reporting mechanisms – report cards, progress reports, transcripts, etc. Those are discussed in Chapter 8.

The PACE accountability option, like competency education as a whole, represents a cultural change that will require the support of all stakeholder groups: students, parents, the community, the school board, teachers, and administrators. When asked about communication during a change process, one curriculum director said, “Start now. Keep going. Don't stop.”

As noted, the specifics of the communication plan will be determined by the district, but some general principles follow:

Start the PACE conversation. Implementing PACE will be a major – and possibly disruptive – change in your district. Begin talking about it with stakeholders as early as possible. Even if all goes as well as it possibly can, it will take time to bring your stakeholders around to supporting this work.

Find allies to lead the PACE conversation. Among all stakeholder groups, there will be people who are ready to embrace the change to a local PACE accountability system. School board members or selectmen/women will be fielding phone calls from community members asking if they know what's going on during a change process. If they can answer, “Yes!” and knowledgeably advocate for the change, then the work of the district will be much easier. Work with these high leverage individuals.

For parents, hearing about the change process from teachers – or, even better, students – is essential. Aside from wanting to know if their child's teacher thinks this focus on performance assessment is a good idea, they also want to know that this teacher is capable of implementing PACE in ways that will help their child thrive. This is very reasonable. The worst thing a stakeholder can hear when asking a teacher about this change is, “I don't really know much about that. You should talk to our principal.”

Engage opponents with respect and compassion. Michael Fullan said, “People don't fear change, they fear loss.” Talk with opponents and try to determine what it is they fear they will lose if PACE is implemented. This puts you in a position address them well. It may even allow you to get beyond their emotional fears and to a place where they can finally hear your arguments. Listen to the opponents instead of shutting them down.

Stage the PACE conversation effectively. The logistics of the conversation can alter the outcome of that conversation. For example, the typical school public meeting, where the

public sits in the auditorium and is addressed from the stage by the principal or superintendent, is a generally poor way to engage stakeholders. An alternative is to hold an event at which food is provided. After the meal, a number of stations are set up at which parents talk with teachers about PACE and performance assessment. Another alternative, similar to “literacy nights,” is to invite parents into the school to discuss PACE and performance assessment with teachers. Another option is for the principal or superintendent to hold “town meetings” or home events to talk with small groups of parents. This is time consuming, but necessary.

The nature of communications through change depends heavily on conditions in the specific district. District coaches will be available to work with district leadership to develop and implement their communications plan.

DRAFT

8. Reporting to the Community

Each PACE district is responsible for annually reporting on student and school progress to parents and the local community, as well as to the New Hampshire Department of Education. These annual reports should include both a district report of aggregate data by grade span (elementary, middle, high) as well as by individual school. The reports should provide a coherent and comprehensive overview across multiple measures, and should capitalize on recent advances in data visualization and graphical representation strategies. Reports should be released at times that school and district leaders find most useful for school improvement planning. While the format of the reports is at the discretion of each PACE district, NHDOE will provide a sample template.

Each district and school report must include the following categories, and for items 2-5, in aggregate form and disaggregated by low-income (free/reduced lunch), race/ethnicity, English language learner, special education, and gender:

1. District/school goals and targets
2. Student demographics
3. Academic achievement and growth

Additional recommended domains include the following:

4. College and career readiness
5. Teacher effectiveness
6. School culture/environment
7. Student engagement in learning
8. Resources

Below are suggested indicators to consider including in each district and school report:

1. District/School Goals and Targets

- Listing of goals, and for each one the yearly targets and the degree to which the target was reached

2. Student Demographics

- Total enrollment
- Enrollment of identifiable student groups

3. Achievement

- Performance assessment results by discipline (English language arts, mathematics, science): percent of students at each level of attainment
- Progress of students within and across years in key content areas
- Smarter Balanced results by discipline (English language arts, mathematics) for one grade in each grade span (elementary, middle, high): percent of students at each level of attainment
- Work-Study Practices: percent of students at each level of attainment

4. College and Career Readiness (high school only)

- Advanced Placement
 - Percent taking AP courses
 - Percent gaining a 3 or better on the exam
- High School Graduation
 - Percent earning one-quarter of graduation credits after one year of high school
 - Percent earning one-half of graduation credits after two years of high school
 - Percent graduating in four years
 - Percent graduating in five years
- Technical Certificates
 - Percent earning professional certificates in a technical center or apprenticeship setting
- Dropouts
 - Percent of entering freshmen who dropped out prior to high school graduation
- SAT
 - Number of students taking the SAT
 - Number of students reaching college-ready benchmarks for mathematics and verbal portions of test
 - Mean scale score
 - Disaggregation of items above by identifiable student groups

5. Teacher Effectiveness

- Total number of faculty
- Percent of educators at each level of attainment (highly effective, effective, needs improvement, ineffective)
- Percent of educators teaching in an area in which they are not certified
- Daily faculty attendance rate
- Annual faculty turnover rate

6. School Culture/Environment

- Results of parent, student, and/or teacher surveys

7. Resources

- Per-pupil expenditure in the current fiscal year
- Average class size
- Student: teacher ratio
- Total computer-to-student ratio

9. District Peer Review Audits: Quality Monitoring of Local PACE Systems

9a. Overview and Purpose

The PACE assessment system will generate considerably more data about student learning than we are used to seeing with tests administered just once each year. However, the PACE system must address the potential lack of comparability among assessment results across schools or even classrooms in addition to concern with significant variability in the quality of the assessments and scoring.

Comparability does not have to come from all students completing the same assessment in the same test window under the same conditions. Instead, “comparability” will result from auditing processes like those used in Queensland, Australia or in Kentucky in the 1990s. In Queensland, for example, school-based assessments used for student accountability go through an extensive, multi-level peer review process to ensure that these assessments meet standards of technical quality and allow students to produce evidence of learning relative to specific curricular targets.

PACE districts will collaborate in a district peer review audit that will assure that districts have in place a strong performance assessment system that meets the criteria of the state-sanctioned PACE option. Participating districts will submit evidence of their performance assessment systems to peer review teams of external practitioners from other PACE districts. The peer review process will be an important vehicle for providing collegial feedback to districts while supporting rich professional learning and cross-district collaboration. The goals of the PACE Peer Review Audit will be as follows:

- 1) Evaluate the quality and components of an effective local PACE accountability system against a shared set of criteria,
- 2) Ensure that PACE districts have in place a strong accountability system that meets key technical quality requirements,
- 3) Provide feedback to districts on the strengths and gaps in their current local assessment and accountability systems and provide recommendations for improving the system, and
- 4) Provide NHDOE with data that can be used to improve the state PACE accountability model.

Districts will be required to submit evidence of their performance assessment systems, according to defined guidelines, to peer review teams of external practitioners, who will review the evidence, gather additional data, and provide feedback. The peer review process is an important vehicle for providing feedback to districts while supporting rich professional learning and cross-district collaboration.

9b. Timeline

The timeline for peer reviews will be staggered. The first two years of reviews will focus solely on providing formative feedback to districts. By the third year, peer review audits will be conducted in order for districts to be approved as a district that has a strong PACE accountability

system in place. Districts undertaking a PACE district review could be approved for up to five years, or receive conditional approval for a shorter time period with specific requirements for improving the local performance assessment system.

9c. PACE Accountability Rubric Categories

The eventual PACE accountability rubrics and the peer review audit will be based on the following categories:

A. Coherent Set of Multiple Measures

- Provides evidence of a coordinated district plan to assess and report on student performance
- Contains measures within, at least, the following domain:
 - Academic (e.g., performance tasks, Work-Study Practices, use of SB in one grade in each grade span)

Additional recommended domains include the following:

 - College and career readiness (e.g., Advanced Placement, graduation, dropout)
 - Teacher effectiveness (e.g., educator evaluation)
 - School culture/environment (e.g., parent, student, teacher surveys)
 - Student engagement in learning (e.g., student surveys)
- Yearly measureable targets are established that are coherent and ambitious

B. System of Performance Tasks to Assess Student Learning and School Progress

- Selection and use of a combination of validated state and local competency-based performance assessment tasks in academic disciplines with state competencies (currently English language arts, mathematics, science) which provide rich data on how students are progressing at every grade span
 - For individual assessment tasks: alignment, fairness, and consistency and accuracy
 - For the system of assessments: alignment, fairness, comparability, and consistency and accuracy
- Process of local task design and scoring that ensures task quality and scoring calibration
- Defines what is acceptable and unacceptable performance for each performance measure
- Local policies that support the state performance assessment system (e.g., performance assessment graduation requirements)

C. Opportunity for All Students

- Must demonstrate the opportunity for an adequate education for all students
- Instructional system is designed to support student learning of competencies
- Academic supports are in place to assist struggling students to meet competencies

9d. Steps in the PACE Accountability Process for FY 2015 and 2016

The PACE Peer Review Audit process will be scaffolded for the first two years as districts build their PACE systems of assessments. For the school years 2014-2015 and 2015-2016, the focus for the PACE District Peer Review Audit will be on evaluating the technical quality of each district's **system of assessments**, *rather than individual tasks*, using the criteria in section 4f. These criteria are all intended to support evaluations of the degree to which competency determinations can be validly made from the assessments:

- ✓ Alignment
- ✓ Fairness
- ✓ Comparability
- ✓ Consistency and Accuracy

In 2014-2015, the process by which we undertake reviewing each district's system of assessments using these criteria will be through three network meetings, to which each PACE district will send district teams:

- **Network Meeting 1: Alignment and Fairness**

Alignment:

- ✓ The set of tasks fully represents the content and skills of the competencies.
- ✓ The set of tasks fully represents the depth of knowledge expected by the competencies.

Fairness:

- ✓ Appropriate accommodations are used so that educators are able to make valid inferences about the achievement of students with disabilities and English language learners.
- ✓ The set of tasks incorporates a variety of task formats so that students who tend to perform better or worse on specific formats are not treated unfairly.
- ✓ Assessment conditions are appropriate for all students.
- ✓ Tasks that reveal differences in mean scores among subgroups are compared to the results on other tasks and external measures (e.g., SBAC and NECAP).

- **Network Meeting 2 – Consistency and Accuracy**

- ✓ The overall generalizability of the set of tasks for a given competency is sufficient for the intended purposes.
- ✓ The system includes carefully constructed scoring rubrics, appropriate training for raters, and a process to monitoring the scoring of student products.

- **Network Meeting 3 – Comparability**

- ✓ If tasks require the same knowledge, skills and abilities and if the scores they yield are equally accurate for students at any given level of those knowledge skills and abilities, then the scores they yield should be comparable.
- ✓ Tasks are related to other measures thought to measure the same or very similar learning targets.
- ✓ The set of tasks produces comparable results across schools and classrooms within the same school district within a given year and across years.

The goals for these meetings will be:

- To provide constructive critique and feedback to each PACE implementing district on their progress toward meeting the technical quality criteria for building a system of assessments
- To provide professional development to participants on building a district system of assessments, while building a professional community of PACE learners

For each network meeting, PACE implementing district teams will be asked to bring evidence to demonstrate their progress toward meeting the technical quality criteria of focus for that day (a list of examples of evidence is listed in section 4f).

The meeting will be structured as a peer review process. Participants will meet in four heterogeneous groups (i.e., representatives from each of the six PACE implementing and planning districts in each group). Within each small group, one of the four PACE implementing district will be assigned to present their evidence for the technical criteria focus to representatives from the other districts. A common protocol will be used in each small group, focused on listening, thoughtful questioning, and informed feedback on the strengths of the district's progress along with recommendations for improvement.

Once completed, the themes from the four districts will be gathered, and common challenges will be discussed as a large group and back in small groups, with debriefing at the end of the day.

Formal notes from each district's presentation and feedback session will be written up and provided to each district to assist in their using the recommendations to further strengthen their system of assessments.

APPENDIX A: District Action Planning Template

District:				
Goal:				
Action Step	Time Line (Completion and landmark dates)	Responsible Person	Resources or Support Needed From NHDOE	Evidence of Completion (How will you know when it is successfully completed?)

APPENDIX B:
Performance Assessment for Competency Education (PACE)
New Hampshire Department of Education
PACE Needs Assessment

The purpose of this assessment is to use as a tool to self-assess your district’s current progress in designing a PACE local accountability system, and determine the additional components that need to be strengthened or developed in order to have a complete system for implementation. This self-assessment will be useful in then completing the PACE district action plan to ensure a comprehensive local PACE accountability system is in place and implemented with fidelity.

Getting Started

Have you:

No.	Question	Yes	No
1.	PACE Team. Established a PACE Team that includes the superintendent, assistant superintendent or curriculum director, special education director, all high school principals, one middle grades principal, and one elementary principal?		
2.	Competencies. Adopted the state K-12 competencies in ELA, math, and science OR adopted local K-12 competencies in these disciplines that are aligned with the state competencies and have been vetted using the Competency Validation Tool?		
3.	Work-Study Practices. Adopted the state Work-Study Practices OR adopted a similar set of local work-study practices that are aligned with the state practices?		
4.	District/School Goals. Set district and school goals for student outcomes (engagement, achievement)?		
5.	Timeline. Created a 2014-2015 timeline of activities/products to be completed in planning or implementing a local PACE accountability system?		
6.	Community Engagement. Engaged community stakeholders (i.e., parents, faculty, students, school board, community organizations) on the shift to PACE?		

PACE Needs Assessment

PACE Performance Assessments

Assess the following statements based on the following criteria:

- 1 = Have not started development
- 2 = In the planning stage
- 3 = Piloting, or testing out with a smaller sample
- 4 = Early implementation with the entire school or district
- 5 = Have revised based on early implementation, and are institutionalizing

Have you:

No.	Performance Assessment Component	1	2	3	4	5	Comments
Performance Measures							
7.	Developed a set of performance measures that are coherent, ambitious, and measurable?						
Professional Development Plan							
8.	Created a faculty professional development plan on designing, administering, and scoring competency-aligned and Work-Study Practices-embedded performance tasks? And that includes:						
9.	✓ Designing validated performance tasks?						
10.	✓ Administering performance tasks?						
11.	✓ Scoring student work consistently across teachers?						
System of Performance Tasks							
12.	Determined a set of performance tasks that will be administered across ELA, math, and science in every grade (inclusive of grades 3-8 and one high school grade)?						
13.	✓ Selected at least one common performance task per content area (ELA, math, science) per grade to administer?						
15.	✓ Vetted local tasks using the QPA validation protocol?						
16.	✓ Aligned local tasks with the adopted competencies?						
17.	✓ Ensured that local tasks are free of bias and can be understood						

	by all students across culture, gender, race/ethnicity, language, and disability?						
--	---	--	--	--	--	--	--

PACE Needs Assessment

No.	Performance Assessment Component	1	2	3	4	5	Comments
18.	✓ Determined that the set of tasks (common and local) fully represents the content and skills of the competencies?						
19.	✓ Determined what acceptable levels of performance are for each task and overall (e.g., for promotion/graduation)?						
Scoring Consistency							
20.	Ensured that the system includes carefully constructed scoring rubrics?						
21.	Provided appropriate calibration training for raters?						
22.	Included a process to monitor the scoring of student products to ensure reliability?						
Overall PACE Accountability System							
23.	Developed and had the school board approve a local PACE accountability policy?						
24.	Aligned PACE with the local educator evaluation system?						
25.	Determined the other measures that will be reported on in the PACE system (e.g., attendance, suspension, graduation, dropout, Advanced Placement, teacher effectiveness rates; school culture)?						
26.	Developed a data system to manage the performance assessment data and other PACE performance indicators and measures?						
27.	Developed a data system that can report out aggregate and disaggregated data by subgroup (race/ethnicity, income, gender, language, disability)?						
28.	Designed a process to use aggregate and disaggregated data to identify student learning gaps and design instructional approaches to address them?						
29.	Designed the reporting system to the community and NHDOE that includes performance assessments & other measures?						

DRAFT

APPENDIX D: WORK-STUDY PRACTICES
New Hampshire Department of Education
Work-Study Practices
Rationale for Work-Study Practices – June 2014

In June 2013, the New Hampshire State Legislature passed Chapter 263, or Senate Bill 48, *An Act Relative to School Performance and Accountability*, which amended the existing statute on school performance and accountability. In this bill, the legislature declared the intent to build a state accountability model that will “best support schools and educators...to enable all students to progress toward college and career readiness with clearly defined learning outcomes.” The legislation underscored the state’s commitment to build a competency-based system in which students are provided with personalized learning that provides “flexibility in the way that credit can be earned and awarded,” and that allows them to “advance when they demonstrate the desired level of mastery.”

As part of this competency-based system, the legislation declared that “New Hampshire’s system of educator support should promote the capacity of educators to deeply engage students in learning rigorous and meaningful knowledge, skills, and *Work-Study Practices* [emphasis added] for success in college, career, and citizenship.” The bill goes on to define Work-Study Practices (WSP) as “those behaviors that enhance learning achievement and promote a positive work ethic such as, but not limited to, listening and following directions, accepting responsibility, staying on task, completing work accurately, managing time wisely, showing initiative, and being cooperative.”

New research highlights the importance of developing and supporting students’ Work-Study Practices. For example, a 2013 United States Department of Education study, *Promoting Grit, Tenacity, and Perseverance: Critical Factors for Success in the 21st Century*, found that non-cognitive abilities (NCAs, their term for Work-Study Practices) “are essential to an individual’s capacity to strive for and succeed at long-term and higher-order goals, and to persist in the face of the array of challenges and obstacles encountered throughout schooling and life.” Hess and Gong, in their paper, “*Ready for College and Career?*” found that “college professors and employers prioritize aptitudes that go beyond typical academic standards, such as communication, collaboration, and creativity” (Hess & Gong, 2014). Angela Duckworth concludes that “grit” or persistence is a better determinant of future success than traditional measures such as a person’s IQ, SAT, and ACT scores (Duckworth & Peterson, 2007). Importantly, research identifies six strategies that can promote New Hampshire’s Work-Study Practices in schools (US ED, 2013; Hess & Gong 2014), alongside the state’s commitment to competency-based education, and that will help lead all students to graduate college and career ready:

1. Focus learning on competencies representing key concepts, skills, and WSPs;
2. Embed WSPs across the curriculum to promote cognitive rigor and deep learning;
3. Provide opportunities for students to take on challenging learning goals that are intrinsically meaningful to them through student-centered learning approaches;
4. Use performance assessments that require demonstration of content, skills, and WSPs;
5. Provide a supportive environment that conveys high expectations and effort over ability;
6. Explicitly teach students how to apply WSPs to their learning, e.g., teaching students how to set goals, define tasks, self-monitor progress, and deal with failures as “bumps in the road.”

**New Hampshire Department of Education
Work-Study Practices and Definitions – June 2014**

In the winter and spring of 2014, a state-wide Work-Study Practices Committee convened, representing teachers and administrators across the state, to develop a set of New Hampshire Work-Study Practices to be approved by the state board of education. The committee defined Work-Study Practices as:

Behavioral qualities or habits of mind that students need to be successful in college, career, and life.

Drawing on research, examples around the country, and New Hampshire legislative language, the committee arrived at a set of four Work-Study practices that all New Hampshire students need to demonstrate proficiency in order to be college and career ready.

Communication

I can use various media to interpret, question, and express knowledge, information, ideas, feelings, and reasoning to create mutual understanding.

Graduating seniors will be able to demonstrate that they can:

- *Communicate effectively using multiple modalities*
- *Interpret information using multiple senses*
- *Demonstrate ownership of the work*

Creativity

I can use original and flexible thinking to communicate my ideas or construct a unique product or solution.

Graduating seniors should be able to demonstrate that they can:

- *Think originally and independently*
- *Take risks*
- *Consider alternate perspectives*
- *Incorporate diverse resources*

Collaboration

I can work in diverse groups to achieve a common goal.

Graduating seniors will be able to demonstrate that they can:

- *Contribute respectfully*
- *Listen and share resources and ideas*
- *Accept and fulfill roles for the purpose of completing a complex task*
- *Exercise flexibility and willingness to compromise*

Self-Direction

I can initiate and manage my learning, and demonstrate a “growth” mindset, through self-awareness, self-motivation, self-control, self-advocacy and adaptability as a reflective learner.

Graduating seniors will be able to demonstrate that they can:

- *Persevere in completing complex, challenging tasks*
- *Use self-reflection to influence work and goals*
- *Engage stakeholders to gain support*

DRAFT

New Hampshire Department of Education Recommendations for Implementing Work-Study Practices

In order for Work-Study Practices to be effective in preparing our students to be college and career ready, schools and districts will need to embed them throughout the curriculum and daily life of students. Following are recommendations for schools and districts to consider in introducing Work-Study practices to your learning communities:

Curriculum and Instruction

- Embed Work-Study Practices across the curriculum K-12, and in all major assignments/projects
- Integrate the language and expectations of Work-Study Practices into daily instruction such as “explicitly teach students how to apply WSPs to their learning, e.g., teaching students how to set goals, define tasks, and self-monitor progress.”

Assessment and Reporting

- Embed Work-Study Practices in all significant performance assessments as a means of assessing students’ WSP proficiency
- Use the sample state-wide Work-Study Practices rubric or create similar local rubrics to assess whether significant assessment tasks are reflective of students exhibiting Work-Study Practices
- Include Work-Study Practices in student report cards and transcripts as a separate assessment

Professional Development

Provide professional development to faculty on embedding Work-Study Practices in the curriculum, assessments, and daily instruction

School Culture

- Post Work-Study Practices widely throughout schools
- Create a school-wide expectation that Work-Study Practices should be evident in students’ daily school interactions, extracurricular activities, etc.

Teacher Evaluation

- Include in teacher evaluation teachers’ integration of Work-Study Practices into curriculum, instruction, and assessment

Parent Engagement

- Educate parents about the importance of Work-Study Practices through newsletters sent home, assemblies, and other means

Annual Assessment

- Annually assess the effectiveness of Work-Study Practices’ implementation and rubrics, and fine-tune based on conclusions

Over the coming months and years, the NHDOE will create a virtual library of resources that can assist educators in local districts to implement the state's Work-Study Practices. These resources will include WSP professional development agendas and materials, sample WSP-embedded curriculum units and accompanying performance tasks, sample report cards and transcripts that include separate WSP assessments, and public relations materials explaining WSP to parents and the community.

DRAFT

**New Hampshire Department of Education
Work-Study Practices Rubric**

	4 - Advanced	3 – Proficient	2 - Improving	1 – Beginning
<i>Communication</i>	Demonstrates all practices in proficient category at a high level and/or consistently over multiple projects	<i>Demonstrates all:</i> <ul style="list-style-type: none"> • Communicates effectively using multiple modalities • Interprets information using multiple senses • Demonstrates ownership of the work 	Demonstrates some but not all practices in proficient category	Demonstrates one or none of the practices in proficient category
<i>Creativity</i>	Demonstrates all practices in proficient category at a high level and/or consistently over multiple projects	<i>Demonstrates all:</i> <ul style="list-style-type: none"> • Thinks originally and independently • Takes risks • Considers alternate perspectives • Incorporates diverse resources 	Demonstrates some but not all practices in proficient category	Demonstrates one or none of the practices in proficient category
<i>Collaboration</i>	Demonstrates all practices in proficient category at a high level and/or consistently over multiple projects	<i>Demonstrates all:</i> <ul style="list-style-type: none"> • Contributes respectfully • Listens and shares resources and ideas • Accepts and fulfills roles • Exercises flexibility and willingness to compromise 	Demonstrates some but not all practices in proficient category	Demonstrates one or none of the practices in proficient category
<i>Self-Direction</i>	Demonstrates all practices in proficient category at a high level and/or consistently over multiple projects	<i>Demonstrates all:</i> <ul style="list-style-type: none"> • Perseveres in completing complex, challenging tasks • Uses self-reflection to complete work and goals • Engages stakeholders to gain support 	Demonstrates some but not all practices in proficient category	Demonstrates one or none of the practices in proficient category

Accountability for College and Career Readiness: Developing a New Paradigm

By Linda Darling-Hammond, Gene Wilhoit, and Linda Pittenger

Why Rethink Accountability?

As states across the country are enacting new college- and career-ready standards, many are seeking to create more aligned systems of assessment and accountability that can assure every child access to the opportunities for deeper learning anticipated by these new standards, and more flexible designs for schools so that their graduates can meet the challenges of a world in which both knowledge and tools for learning are changing rapidly.

While the evolution of federal policy has contributed to advances over the last two decades — in particular, the focus on learning standards begun in the Clinton administration and the expectation that “every child counts” under No Child Left Behind (NCLB) — it is clear that these prior efforts are inadequate to the current challenges.

Although gains have been registered on the state tests that have been the focus of accountability under NCLB, U.S. performance declined between 2000 and 2012 on all subjects in the Program for International Student Assessment (PISA) — a more open-ended set of assessments evaluating how students can apply their knowledge and solve problems. On all of these measures, large and persistent achievement gaps remain among U.S. students by income, language background, and racial and ethnic group. The United States also exhibits one of the highest rates of childhood poverty in the developed world while distributing far fewer of its educational resources to meet the needs of disadvantaged students.

If we want to ensure that all students are indeed prepared for college and career readiness in these needs, several major changes are required. Among them are:

- More sophisticated curriculum and assessments “of, for, and, as learning” that foster and evaluate deep understanding of content, critical and creative thinking, problem solving, multiple modes of communication, and uses of new technologies to find, synthesize, evaluate, and use information to answer questions and create new solutions.
- More equitable and adequate resources which ensure that all students have access to the quality of teaching, materials, and technology they need to engage the new standards productively, and which address the additional needs of students who live in poverty, are new English learners, or who have other special educational needs.

About New Accountability

New standards require major changes in curriculum, assessment, and school organization that, in turn, require new forms of accountability. If educational improvement is the goal, these should focus on meaningful learning, professional capacity, and adequate resources, wisely used.

Stanford Center for
Opportunity Policy
in Education

650.725.8600
edpolicy.stanford.edu
@scope_stanford

scope
Stanford Center for
Opportunity Policy in Education

cie national center for
innovation in education
University of Kentucky

- Greater capacity among schools and educators to teach this more challenging content to an increasingly diverse group of students. This will mean developing pedagogies for deeper learning focused on 21st century competencies, personalizing instruction, and creating school designs that allow students to learn and apply their knowledge in ways that take advantage of new technologies and link to the world beyond traditional school walls.
- A more effective model for change and improvement that can foster the collaborative changes needed to transform schools from the industrial model of the past to innovative learning systems for the future. Rather than placing schools in a straitjacket reflecting the demands of tests pointed toward the past, accountability will need to enable thoughtful risk-taking informed by continuous evaluation using multiple measures to inform improvement.

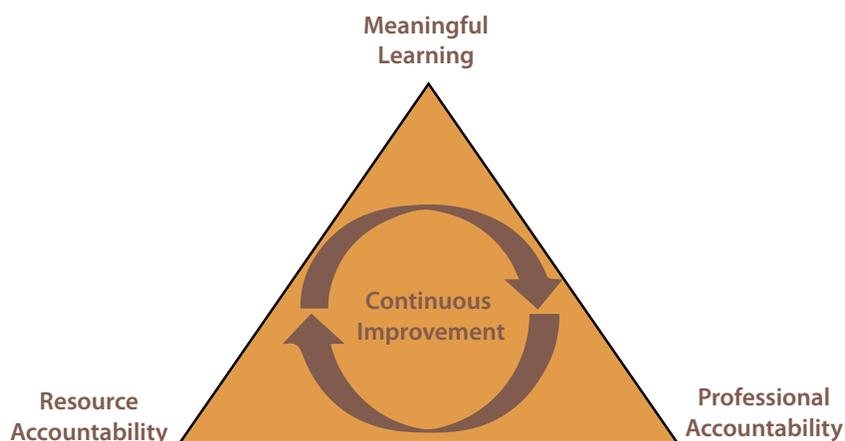
What Should a New Approach to Accountability Entail?

Since 2002, federally-enforced educational accountability has been defined primarily as

the application of specific consequences to schools that do not meet annual targets for growth on yearly state tests. More is needed to meet current demands, however. If the goal of an accountability system is to improve education, it must raise expectations not only for individual schools but for the functioning of the system as a whole — and trigger the intelligent investments and change strategies that make it possible to achieve these expectations. This should include well-articulated expectations for what states and districts should do to provide the resources or conditions for learning, along with well-developed systems for improving professional skills, and research-based processes for guiding change and improvement.

A good starting point is to consider what parents and the public need an education system to be held accountable for: that children be taught relevant and meaningful skills that will prepare them for the world they are entering and that they be taught by competent professionals in adequately resourced schools responsive to their needs. From this perspective, a new paradigm for accountability should rest on three pillars: a focus on meaningful learning, enabled by professionally accountable educators, supported by adequate resources that are well-used. It

Key Elements of an Accountability System



should be animated by processes for continuous evaluation and improvement that lead to problem solving and corrective action at the local level.

In such a system, accountability should be:

Reciprocal: Each level of the system — from federal and state governments to districts and schools — should be accountable for the contributions it must make to produce high-quality learning opportunities for each and every child. States and districts must be accountable for providing the resources, supports, and incentives that result in well-staffed, effective schools. Schools must be accountable for using these resources wisely and enabling strong teaching. Educators must be accountable for teaching the standards in ways that respond to their students' needs. Everyone must be accountable for continuous learning.

Focused on capacity-building: An accountable system acts on what is known about best practices: It builds capacity by making knowledge about what works widely available and provides learning opportunities for practitioners and policymakers, so that this knowledge is well-used.

Committed to problem-solving and improvement: An accountable system creates and shares transparent data and information, along with strategic evaluation processes, like school quality reviews, that can identify problems and guide diagnosis and corrective action.

Accountability Should Focus on Meaningful Learning

If meaningful learning for all students is the focus of an accountability system then cur-

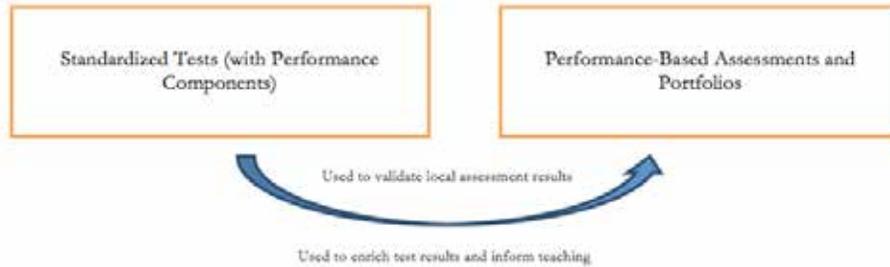
riculum, assessment, and instruction must support the knowledge, skills, and dispositions students will need to succeed in 21st century college, careers, and citizenship — including the abilities to solve problems and apply knowledge, inquire and learn independently, build relationships, use feedback, and persevere in the face of obstacles.

Capturing and supporting meaningful learning will require richer assessments that more authentically evaluate 21st century skills. These should be used to inform teaching and to expand, rather than limit, educational opportunities for students.

A System of Higher Quality Assessments: Assessments, both state- and locally-administered, should include more open-ended items on summative tests, along with classroom-embedded performance tasks — research inquiries, scientific investigations, literary analyses, mathematical models, written and oral presentations, technology products — that develop and assess higher order skills. Robust performance assessments can also support and evaluate harder-to-measure abilities that matter greatly to success: the abilities to collaborate; to plan and organize time, materials, and people; to overcome obstacles; to persevere; to use feedback productively; and to learn independently.

New York State, for example, has authorized schools in the New York Performance Standards Consortium to use a portfolio of performance assessments with common rubrics and scoring, in lieu of the Regents tests in most subject areas. Envision Schools and many Linked Learning schools in California use a similar approach. Research has shown that graduates from these networks of schools have higher college-going and col-

A New Competency-Based System of Assessment



lege success rates than other students in their states¹.

New Hampshire is currently implementing a plan for a new competency-based system relying on a combination of state and local performance assessments to supplement the Smarter Balanced Consortium tests based on this design. The state will use a smaller number of higher-quality state tests to validate local judgments based on evidence from more in-depth tests and tasks, which offer more detailed information about how students think and perform, and can guide more effective teaching. This new system of assessment will move from an overemphasis on external summative tests to a greater emphasis on performance assessments that can inform and improve learning.

Accountability Should Ensure Adequate Resources, Wisely Used

In a country where school funding inequities are severe, inadequate resources deny genuine accountability to many families. If

Footnote: Foote, M. (2007). Keeping accountability systems accountable. *Phi Beta Kappan*, 88(5), 359-363; New York Performance Standards Consortium. (2014). *Educating for the 21st century: Data report on the New York Performance Standards Consortium*. http://performanceassessment.org/articles/DataReport_NY_PSC.pdf; Friedlaender, D., Burns, D., Lewis-Charp, H., Cook-Harvey, C. M., Zheng, X., & Darling-Hammond, L. (2014). *Student-centered schools: Closing the opportunity gap*. Stanford, CA: Stanford Center for Opportunity Policy in Education.

we really expect all children to achieve college and career-readiness, governments at all levels must be accountable for fairly allocating and wisely using resources — dollars, curriculum and learning tools, well-qualified educators, time, and safe, healthy environments for learning — to accomplish these goals. Measures of resource adequacy must become part of the accountability system, along with indicators of system performance that allow the public to understand what is being invested and with what results.

Resource Standards: Allocating adequate resources in relation to students' learning needs should include ensuring equitable access to high-quality curriculum and instructional materials that support students in learning the standards; providing well prepared educators and other professional staff to all students in settings that allow them to attend effectively to student needs; and ensuring additional supports for students with particular needs associated with poverty or educational requirements.

Transparency: Data and information should be made available to the public on how funds are spent and what outcomes result. This is a key aspect of the accountability strategy to support analysis of resource use.

Multiple Measures: To evaluate whether resources are adequate and appropriately used, multiple measures of access and performance for students, educators, and

schools are needed to inform decision making at each level. These should capture the multi-faceted aspects of education valued by parents, the profession and community. Like the dashboard on a car, which provides indicators of speed, distance traveled, fuel, fluids, tire pressure, and more, the combination of measures signals where to look further to figure out how things are working.

California’s recently adopted Local Control Funding Formula (LCFF) is an example of an approach that addresses all of these elements. The LCFF allocates all funds based on pupil needs (weights are applied to pupil counts based on poverty, English learner status, and foster care status). The accompanying Local Control Accountability Plan requires districts to develop, adopt, and annually update — with parent and community involvement — a 3-year accountability plan that identifies goals and measures progress across multiple indicators of both opportunities and outcomes. Local districts can add their own indicators to those that are state required. Data are *disaggregated* by student race and ethnicity, poverty, language status, and disability status. Indicators must include:

- Student achievement: State tests and other assessments (e.g., AP or IB tests, English proficiency)
- Student persistence and graduation
- Student inclusion (suspension and expulsion rates)
- College- and career-readiness indicators (access to and completion of curriculum pathways)
- The availability of qualified teachers, adequate facilities, and necessary materials
- Student access to a broad curriculum, including the core subjects (including science and technology), the arts, and physical education
- Evidence of parent participation and opportunities for input

Districts can add to the state measures, as the set of seven California Office to Reform Education (CORE) districts (Fresno, Long Beach, Los Angeles, Oakland, San Francisco, Santa Ana, and Sanger) did in their federal Elementary and Secondary Education Act (ESEA) flexibility waiver, when they added evidence of social-emotional learning and school climate, for example. Surveys of teachers, parents, and students are part of the data that help schools become more aware and responsive. The CORE accountability structure is shown below.

CORE Accountability Structure



Problem-Solving and Corrective Action: These data should be evaluated through well-designed systems of review, judgment, and intervention, rather than being used mechanically to mete out sanctions.

Such systems — whether evaluating student learning, educator performance, or school performance — should involve experts in interpreting information to guide consequences or corrective action based on a deep understanding of what is happening and what is needed. The goal should be to make strategic changes that protect students' rights and promote system improvement.

Accountability Should Support Professional Capacity and Ensure Competence

Unless students experience good teaching, accountability is meaningless. Accountability for implementing professional practice rests both with individual educators and with the schools, districts, and state agencies that recruit, train, hire, assign, support, and evaluate staff and organize education. Collectively, they are responsible for ensuring that the best available knowledge about curriculum, teaching, assessment, and student support will be acquired and used by individual educators and by the system as a whole.

The heart of a professional accountability system is a set of elements that ensures that educators are carefully selected, receive high-quality preparation that enables them to acquire essential knowledge and skills, are licensed based on useful evidence of effectiveness, supported through high-quality induction and professional learning opportunities, and make sound personnel decisions — including opportunities for advancement that support further sharing of expertise — through thoughtful evaluation, supervision,

and career ladders. Professionally accountable systems also ensure that well-qualified educators are readily available to all students across the state, which requires attention to recruitment incentives such as service scholarships, adequate and equitable salaries, and working conditions that provide motivation to stay.

Professional standards of practice should guide how educators are prepared and how they teach, lead, organize schools, and support students. States should adopt and use professional standards aligned to student learning standards to guide preparation, accreditation, licensure, and practice and to build capacity at all levels of the system, including:

- Educator capacity that enables teachers to teach for deeper learning and administrators to understand and support this work at the school and district level. This requires:
 - High-quality preparation, induction, and professional development
 - Accreditation and licensing based on evidence of teacher and administrator performance in supporting diverse learners to meet challenging standards
 - Evaluation based on multiple indicators of practice, contributions to student learning, and contributions to colleagues in support of student learning and schoolwide improvement.
- School capacity to meet student needs, based on school, district, and state actions that ensure the availability of an appropriate mix of well-qualified staff who are properly assigned and adequately supported with professional development, and who are engaged in well-de-

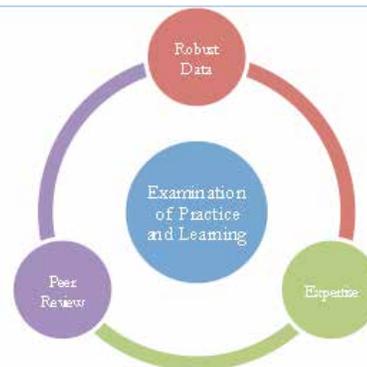
signed curricula and educational programs that are consistent with research

- System capacity for professional practice and improvement, supported by awareness of research, as well as inspection or school quality review processes, that evaluate policies, programs, practices, and outcomes, diagnose areas for improvement, and guide appropriate interventions.

A School Quality Review system should help schools assess their practices and work on areas for improvement, supporting well-guided intervention and corrective action in schools where the evidence suggests that achievement is not adequate and students' needs are not being met.

An effective School Quality Review process should bring together several elements that have not been joined before in most education policy systems: robust data, educational expertise, and peer review. Like the inspectorate model used in many countries, it should be guided by experts who are deeply knowledgeable about practice and well-trained in how to conduct a diagnostic inquiry into school practices and their relationship to the nature and quality of student learning. Like U.S. accreditation systems, the engagement of peer reviewers from other schools in the state can enlist multiple perspectives while stimulating a learning process that expands the knowledge and sharpens the analytical skills of participants. Like many research endeavors, the skillful use of robust quantitative data, much of which is comparable across schools, with qualitative insights developed from looking purposefully at teaching and student work and talking to stakeholders, can allow reviewers to get a better understanding of how the school is working and what may help it improve. By combining these things, such a process is

Elements of a School Quality Review



more powerful and purposeful than accreditation approaches have been in the past.

School quality review approaches like this have been used successfully at various times in Kentucky, New York, North Carolina, and Rhode Island. Teams of distinguished educators are typically then called in to support the hands-on work of school improvement based on the deep analysis that has been provided. In some cases, these efforts have been focused on struggling schools. They are able to reveal what it will take to improve a school; whether changes are needed in curriculum, leadership, staffing, or other aspects of the organization; and even whether students would be better served by closing and redesigning a school entirely. While struggling schools or districts may engage more intensively in such reviews and follow up efforts, a School Quality Review process should ideally be used to support system-wide learning and improvement.

Similarly, peer assistance and review programs have been used successfully in teacher evaluation to bring the expertise of mentors and the judgments of a panel of teachers and administrators to bear for helping teachers to improve, and making decisions about removal where improvement does not follow intensive assistance. In both cases, adding expertise, peer evaluation, and carefully

collected data to a process of review and assistance around standards of professional practice produces better-grounded analyses and more effective decisions.

Engaging teachers in jointly scoring student work and consulting about how to improve curriculum and teaching to produce greater success for learners also helps build professional norms and knowledge. Indeed, engaging students in reviewing their own and their peers' work to guide revisions in light of standards leverages powerful learning. Professional capacity and accountability are reinforced by systems of professional judgment for evaluating the work of students, teachers, and schools. Not only does expert professional judgment — used to make sense of qualitative and quantitative information — support more defensible decision, it can also help professionalize education by supporting educators' learning and sense of responsibility as they work with students and families to engage in accountability themselves.

Conclusion

We believe that a new conception of accountability can help the nation meet its aspirations for preparing college- and career-ready students by:

- addressing the opportunity gap that has allowed inequalities in resources to deprive many students of necessary opportunities to learn;
- developing curriculum and assessments that are focused on 21st century learning skills and used in ways that support improvement in teaching and learning;
- creating a dashboard of multiple measures to evaluate schools and sophisticated strategies, including school quality reviews, for helping them improve;
- developing professional capacity, through

high-quality preparation, professional development, evaluation, and career advancement for individuals, plus sharing of expertise within and across schools.

One account of what this new accountability model would look like in a state that developed an integrated system can be seen at [**\[list url\]**](#).

The gauge of a new system should be the outcomes it enables. True accountability should allow schools to be both responsible for high-quality professional practice and responsive to students' needs within the context of their families and communities. An effective accountability system should give students, parents, and governments confidence that schools are focused on what matters most and capable of helping each child connect to a productive future.

Acknowledgments

This brief is derived from a report developed in the course of a series of discussions about the design of a new accountability system, convened by the Hewlett Foundation. We are grateful to the Foundation for its support of this work. The participants in these discussions offered substantial input and ideas. There was strong agreement about many of the ideas, as well as diverse perspectives about some. The final product was authored by Linda Darling-Hammond, Gene Wilhoit, and Linda Pittenger. It reflects the individual and collective insights of the participants, but it does not reflect an endorsement by any of these individuals or the organizations with which they are affiliated. These intellectual contributors include:

Stephen Bowen, Council for Chief State School Officers
Anthony Bryk, Carnegie Foundation for the Advancement of Teaching
Richard Carranza, San Francisco Unified School District
Michael Cohen, Achieve
Linda Darling-Hammond, Stanford Center for Opportunity Policy in Education
Michael Kirst, California State Board of Education
Paul Leather, New Hampshire Department of Education
Philip Lovell, Alliance for Excellent Education
Carmel Martin, Center for American Progress
Jal Mehta, Harvard University
Charmaine Mercer, Alliance for Excellent Education
Rick Miller, Center for Organizational Reform of Education
Scott Palmer, Education Counsel
Linda Pittenger, National Center for Innovation in Education, University of Kentucky
Arun Ramanathan, Education Trust West
Larry Rosenstock, High Tech High School
Gene Wilhoit, National Center for Innovation in Education, University of Kentucky

*Stanford Center for
Opportunity Policy
in Education*

Barnum Center
505 Lasuen Mall
Stanford, CA 94305
scope@stanford.edu
650.725.8600
edpolicy.stanford.edu
@scope_stanford

scope
Stanford Center for
Opportunity Policy in Education

c!e national center for
innovation in education
UNIVERSITY OF KENTUCKY