

The New England Common Assessment Program: Notes on the Collaboration Among Four New England States

**compiled by Charles A. DePascale
Center for Assessment (NCIEA)**

**on behalf of the NECAP states:
Maine, New Hampshire, Rhode Island, and Vermont**

November 2009

Educational and economic factors have intersected to make the concept of collaboration among states to produce common standards and common assessments a high-profile topic at the national level as well as at the regional level. Eight years into No Child Left Behind, there is a growing belief that the disparity in performance across states (i.e., percentage of proficient students) can best be addressed through common standards and common assessments. Current economic conditions combined with concerns about international competitiveness in a global economy also make common standards and common assessments attractive approaches to save money in the short-term and produce students better prepared to compete globally in the long-term.

Recent frenetic activity at the national and federal levels, particularly with regard to the Common Core State Standards Initiative (CCSSI) sponsored by CCSSO and NGA, the USED \$350 million Race to the Top assessment grant, and the possible interrelationship between those two projects have raised interest in issues related to cross-state collaboration and assessment consortia to a new level.

Currently, four New England states Maine, New Hampshire, Rhode Island, and Vermont serve as a national model of the positive potential of collaboration among states. The high-profile New England Common Assessment Program (NECAP) has drawn national interest since its inception in 2004. The success of Rhode Island, New Hampshire, and Vermont on the recently released NAEP 2009 Mathematics tests is being touted as evidence of the efficacy of and need for common standards and assessments.

Introduction: Desperation, Innovation, Collaboration

The relationship among desperation, innovation, and collaboration is well-established, and cannot be underestimated in the case of the collaboration among states that has resulted in the New England Common Assessment Program. In 2001, states were faced with increasing their statewide testing programs from three grades to seven grades to meet the requirements of impending No Child Left Behind Act (NCLB). New Hampshire, Rhode Island, and Vermont concluded that it would be impossible for them to maintain the type of state assessment program they desired across seven grade levels.

Like many states, during the 1990s New Hampshire, Rhode Island, and Vermont implemented state assessment programs that went beyond traditional, norm-referenced, standardized, multiple-choice tests. Rhode Island and Vermont adopted the standards

and assessments of the New Standards Project as their state assessment. Although perhaps falling short of the project's original goals, the New Standards exams available in 2002 still included broad-based tasks designed to measure problem solving and other high level skills. The states were also involved in landmark efforts in portfolio assessment and direct writing assessment. In the early 1990s, New Hampshire began development of the New Hampshire Educational Improvement and Assessment Program (NHEIAP) which included a custom-developed state assessment at grades 3, 6, and 10. In addition to constructed-response items, the NHEIAP tests also included innovations such as a video-based listening test.

Even with additional federal funding provided under NCLB, the states were left with two options in 2002-2003:

1. Switch to primarily multiple-choice tests that were commercially available.
2. Explore options for collaboration with other states that would allow them to continue to offer the type of assessments they valued.

Developing and administering custom state assessments as individual states was not an option. With no attractive, alternative options available the will to succeed and the commitment to make the collaboration work was strong within each of the states.

It is worth noting that New Hampshire, Rhode Island, and Vermont were not alone in initial discussion regarding a collaborative effort among New England states. Preliminary discussions included all six New England states. Connecticut and Massachusetts, with custom state assessment programs and the ability to scale to seven grades to meet NCLB requirements were not in the same position as New Hampshire, Rhode Island, and Vermont and quickly left the conversation. Maine remained in the conversation through the initial development of common content standards but left the discussion of common assessments in an attempt to preserve the well-established Maine Educational Assessment (MEA) at grades 4, 8, and 11. After several cost-saving adjustments to the MEA across five years, Maine ultimately was faced with the same "multiple-choice or collaboration" decision as the other states. In November 2008, Maine became the fourth partner state participating in the New England Common Assessment Program.

Level of Collaboration, Control, and Commitment

After deciding to collaborate on a common assessment, the states were faced with a series of decisions regarding the level of collaboration that they intended, the amount of control they expected to exercise with regard to the development and administration of the common assessment, and the level of commitment each partner state would make to the consortium. The initial meetings among the states to discuss and resolve these issues played a major role in defining the partnership and establishing the principles that have guided the operation of the consortium during the last six years. The investment of start-up time is essential to building clear expectations, common language, strong lines of communication, and partner trust, as well as indentifying the differing strengths among the states that can be used to meet the needs of the consortium as a whole.

Level of Collaboration

Having agreed to collaborate on a common assessment program, the states had to define what was and what was not included in that program – what’s on the table. At an initial meeting, the states decided that their accountability systems and alternate assessment programs (AA-AAS) would not be included in the collaboration. It was also decided that this collaboration would not impinge on individual data warehouse projects that each state had already started. In part, this decision reflected and respected the differences in approaches and philosophies the states had in these areas (e.g., percent proficient v. proficiency index as the basis for accountability, level of emphasis on progress v. status on the alternate assessment). In part, however, the decision also reflected recognition on the part of the states of the need to limit their task to areas directly related to producing a series of operational tests for the 2005-2006 school year. In that same spirit, the states decided to delay collaboration on a common high school assessment and to defer a decision on collaboration on a common science assessment¹. The consortium agreed that the initial collaborative assessment effort would be limited to grade 3 through 8 tests in reading, mathematics, and writing (grades 5 and 8) – not coincidentally, the areas in which the states’ collaborative development and dissemination of common content standards was well underway.

At this point in the process, the states also reached agreement on what would be *common* in their common assessment program. Prior to often intense discussions surrounding specific details, it was not always clear which aspects of the program needed to be common, which components it was desirable to have in common, and which could be kept unique across the states. Critical, high-level, components that would be common among the states included:

- All test materials (e.g., test booklets, answer documents, manuals) would be common and would include the program name rather than individual state names.
- There would be a single set of achievement standards across the states.
- There would be common administration procedures across the states.
- There would be a common policy on the use of accommodations and a single table of allowable standard test accommodations².
- There would be a single set of reports – although compromise led to an expansion of what is included in the single set of reports and variation in the emphasis/use of the particular reports across the partner states.

¹ Development of common high school standards began immediately and a common high school assessment was administered operationally during the 2007-2008 school year. Joint meetings among the states’ science specialists to review existing content standards also led to the decision to collaborate on a common assessment program at grades 4, 8, and 11 that also had its first operational administration during the 2007-2008 school year.

² Within the common table of Standard Test Accommodations, there was an option for schools to request state DOE approval for accommodations not included on the list. Approval of these requests was granted at the state level with the understanding that state decisions would not contradict consortium policy explicitly or implicitly expressed through the table of Standard Test Accommodations.

- There would be a common administration period – although the period became slightly expanded to accommodate the variety of locally observed holidays, teacher workshop days, etc. across the three states.

There was also agreement on some key components regarding the program that would not be common:

- Although there would be a single contractor, there would be individual contracts with each state rather than a single contract across states³.
- Although all technical analyses such as item calibration, scaling, equating, standard setting would be conducted with data combined across the states, there would be no reporting of combined NECAP results. Results would only be reported at the individual state level.
- The purpose of the program was not to make comparisons across states, and there would be no reports directly comparing the results across states.
- The release of results would be handled individually by each state.

Finally, there was also agreement on the general principle that the states would continue to work together to keep as much common across the consortium as possible.

Level of Control

A key decision that the states had to make as a consortium was the level of ownership and control that they wished to have over the design, development, and implementation of the common assessment program. Rhode Island and Vermont were part of the New Standard Project – a collaborative in which they were involved in some aspects of the development of the tests but exerted little direct control over the program and did not own the assessments. New Hampshire had a custom-developed state assessment in which working with their assessment contractor they were deeply involved in almost all aspects of the program and owned all materials developed for the program.

The states determined that for the purposes of this assessment program they would be working with an assessment contractor to design, develop, and administer a set of custom assessments that would be jointly owned by the states⁴. In many respects, the program would operate in the same manner as a single state implementing a custom state assessment. It should be noted that the size of the consortium, the states, and departments of education in this collaborative played a role in the execution of this decision. With the consortium consisting of three (and now four) small states with small numbers of assessment staff within each state the size of the program and the total number of state staff involved in the process was comparable to some single-state assessment efforts. The physical proximity of the consortium states and the contractor also played a key role in the execution of this decision.

³ The task of determining how to enter into a single multi-state contract was too daunting and time-consuming for the New England consortium.

⁴ Like many aspects of this consortium, a legal definition of joint ownership has not been established.

Level of Commitment

From the very beginning of the collaboration, the states had to make a decision on the level of commitment that they were willing to make to the project. Beyond their commitment to the success of the project, this decision involved the commitment of staff time and related resources to the ongoing operation of the project. As discussed in the previous sections, the decision to form a consortium and collaborate on a common assessment program can require various levels of effort and responsibility on the part of the partner states.

The states determined that the assessment and content staff from each of the consortium states would be fully involved in the administration of the New England Common Assessment Program – consistent with their decision to treat this project as a custom state assessment. This decision required the ongoing involvement of the state assessment directors in the management of NECAP and the ongoing involvement of the state content specialists in the test development process throughout the test development cycle. From the beginning, this decision placed additional, significant burdens on the state staff in terms of the time and effort needed to reach agreement with their counterparts in other states as well as the additional physical demands of meeting with counterparts in departments across states rather than across offices within a department.

Although cost was a key factor influencing the states decision to collaborate to attain the quality assessments that they desired, available staff resources was also a critical factor. With their existing staff, it was not feasible for the states to consider developing the type of custom assessments that they desired at grades 3 through 8 and high school in reading and mathematics – as well as additional assessments in writing and science. There was an expectation that collaboration across states would lead to an overall increase in the number of state staff available to participate in the process, and that the act of working together with colleagues from other states to discuss issues related to the assessment development would result in a better product than could be produced by a single state even with the same number of staff (i.e., the whole is greater than the sum of its parts). For the collaboration to achieve this expectation and be maximally effective, however, a division of labor among states is essential. This division of labor (e.g., allowing other states to take the lead on particular tasks), however, requires a fair degree of trust among participants that can only be developed over time and fostered through careful planning, facilitation, and the development of a common understanding of which aspects of the project require full review and approval by all states.

Layers of Support

One lesson learned through the initial work on the development of common standards, subsequent work over the last six years on the common assessment program, and additional collaborative efforts that have emerged during that period (e.g., Enhanced Assessment Grants, WIDA, secondary school reform/restructuring efforts) is that layers of support from within and outside of the departments of education are needed to make the project successful. Within each state, support from the legislature, governor's office, board of education, commissioner, and critical units within the Department of Education is needed to implement a collaborative project that requires giving up some local control

and policies and practices in support of the common goals. Maintaining bottom-up design principles and local involvement/buy-in in a regional, collaborative effort that by its nature tends to favor a top-down rather than bottom-up approach also requires tremendous support. Finally, finding the appropriate external partners to support each phase of the projects has been a critical factor in their success. In addition to specific content expertise, the importance of external support in organizing, facilitating, and negotiating among the partner states cannot be underestimated.

As a matter of best practice, and as required by law, state boards of education (and to some extent legislatures and Governors) are directly involved in the adoption of state content and achievement standards as well as in policy shaping the design, purposes, and uses of the state assessment. In a collaborative program such as NECAP, those groups must join the assessment staff from the Department of Education in their willingness to cede some degree of individual control over issues such as the development of achievement standards and the establishment of cut scores on the individual tests. Although there are real and perceived benefits of collaboration that are readily obvious to policy makers within each state, it is necessary to devote sufficient time and resources to developing an understanding of the impact of collaboration on local control. The policy makers also have an understanding of smaller issues such as the need to remain supportive of cross-state efforts at times when out-of-state travel is curtailed or eliminated as a routine cost-savings measure in an historically insular economic and political environment such as a state.

At least as important as the commitment to the project of state policy makers is the support, buy-in, and participation of local educators and other stakeholders in the project. One of the key lessons learned during the last two decades of development and implementation of high quality content standards and standards-based assessments is that local educators are critical to the successful implementation of standards and assessments. The importance of significant stakeholder involvement in development and implementation of content standards and the importance of efforts to disseminate those content standards have been documented across states. With regard to assessment, each of the partner states involved in NECAP has long histories of teacher participation in the assessment program through membership on committees, participation in scoring institutes, and related professional development activities. The level of understanding and acceptance of the assessment program that results from these efforts is substantial.

Participation in a consortium limits the availability and feasibility of many traditional opportunities for teacher involvement in the assessment process. For example, at the state level item review committees often involve the participation of 20-25 teachers per committee. With multiple committees and teachers participating on a rotating basis, several hundred teachers and schools across the state are directly participating in the assessment program within a relatively short period of time. Scoring institutes in which teachers participate in the operational scoring of assessments has been another common activity in some of the states – particularly in the case of the writing assessment. Operating within a consortium, however, given that the number and size of the committees cannot simply be increased, the number of participants per state must be

reduced⁵. Consequently, the number of educators actively participating in the assessment program within a year and across years is reduced, and the exposure and understanding of the program across the state is diminished. It then becomes necessary for the states to develop and offer other avenues for local educators to participate in the assessment.

Managing the Collaboration

Another lesson learned through the last eight years is that although collaboration among several New England states to develop a custom common assessment program has many benefits, the mere consideration of collaboration is a process that requires planning, organization, and a great commitment from each of the partner states. Further, the shift from a conceptual discussion of collaboration to an operational implementation of a program such as NECAP requires an exponential increase in planning, organization, and commitment along with the additional element of compromise. To a large extent, the success or failure of a multi-state collaborative is dependent on the up-front decision that are made on how the project will govern itself and the processes and procedures that are established to manage the project. From their initial efforts in the development of common standards through their shared management of NECAP, the partner states in the consortium have incorporated external support from organizations such as Education Development Center, Inc. (EDC) and the Center for Assessment (NCIEA) into the management of the consortium.

During the process of developing the initial assessment RFP, the states determined that it would be necessary for them to have someone serving in a program management role to coordinate the efforts of the states. The primary function of this person would be to coordinate the efforts among the states and to serve as a liaison between the states and their assessment contractor. Note that this is different from the project management role served by staff of the assessment contractor. A key part of the states' program management effort is reaching consensus on major program issues before placing demands on the assessment contractor. The states considered multiple options for this management function including a) on a rotating basis, assigning a single state assessment director or staff person as project director for a specified period, and b) hiring a permanent employee as consortium project director through one of the state departments of education, before deciding to contract with an external organization to provide management services in support of the consortium.

As the process of designing, developing, and implementing a common assessment program began, it quickly became clear that several procedural and governance decisions would have to be made. One of the first decisions was whether the consortium would exist as a separate legal entity or as loosely coupled association of individual state departments of education working together for a common purpose. Having decided to move ahead as an informal federation of states committed to the success of a common assessment project, the consortium states had to make a series of key procedural decisions that would shape the consortium, including:

⁵ As the size of the consortium increases, participation on some committees may be limited to state department staff or a single local educator.

- Will there be some decisions that require unanimous agreement and others based on majority opinion?
- On issues where there is not a clear consensus among the states, how will each state have an equal vote in making the decision or will voting be proportional based on a factor such as K-12 enrollment? Although New Hampshire has twice as many students as Vermont, the relatively small and similar size of each of the states helped make the desired solution of one state/one vote feasible. That decision may have been more difficult with the addition of a state such as Massachusetts where the combined K-12 enrollment of the four NECAP states is only 70 percent of the K-12 enrollment of Massachusetts or a state such as California with a K-12 enrollment more than 10 times greater than the combined K-12 enrollment of the NECAP states.
- How will fiscal and contractual requirements be handled? Which costs will be shared equally and which costs will be divided proportionally?

Summary

Beyond the formal procedures and rules established for governance, however, the operation of the consortium has been based on a shared sense of purpose of coming together toward a common cause and a sense of trust that an individual state's willingness to compromise on one issue will be met by a similar willingness by the other states on subsequent issues – without the need for formal quid pro quo agreements among all of the states, or worse yet, subgroups of states. There has also been an understanding of the importance of not requiring each state to abandon all of its unique practices and traditions. Of course, there have been cases where different opinions or philosophies of assessment, instruction, or a particular discipline, in general, make it clear that a debate on an issue is not going to end with a clear consensus. For those cases, norms have been established to end debates and either make a decision or, in the case of content committees, defer a decision to the management team.

The discussion presented in this document has focused almost exclusively on the role of the states in the operation of an assessment consortium. Before concluding, it is important to note that the assessment contractor(s) working with the consortium also play a critical role in determining the success or failure of the consortium. The contractor must be strong and flexible enough to engage in the work of the consortium with an understanding of which areas of the assessment program require complete uniformity across states and those areas in which the consortium might be strengthened by slight variations across states that do not damage the quality of the product or impact the overall cost of the project. It can require a substantially different set of skills to function as a general contractor serving multiple states on a common project than to function as a vendor selling the same product to multiple, individual states.

As a final point, it is also important to note that the experiences of the states in the management and operation of NECAP are shaped by the particular context of that assessment – that is, the NCLB requirements, size of the program, physical proximity of the states, etc. The NECAP states are also active participants in the WIDA consortium and administer the ACCESS for ELLs tests of English language proficiency. The

management and operation of that consortium has some similarities to and some differences from the NECAP consortium that are better suited to its context. It would be a mistake to conclude that all consortia must operate in a similar manner or that a state cannot effectively participate in multiple consortia.