

# Establishing a Common Core Assessment Program: A Contract Conundrum

---

Charles A. DePascale<sup>1</sup>  
National Center for Improvement of Educational Assessment  
December 2013

*States attempting to adopt a common core aligned state assessment, particularly those committed to a national consortium, are finding the procurement process to be more complicated than in the past. In particular, efforts to identify an assessment prior to the RFP process are regarded as awarding “sole source” contracts outside of the intent and letter of state procurement policies. Understanding the ways in which the emerging common core assessment landscape is similar to the current assessment development process and how it differs can be helpful in crafting an appropriate argument to successfully navigate through the procurement process.*

---

States planning a smooth transition from their current state assessments to the assessments under development by one of the national consortia, Smarter Balanced or PARCC, are facing a contract conundrum. Simply stated, how does a state accomplish the task of adopting PARCC or Smarter Balanced as its state assessment program while a) meeting the requirements of its state procurement regulations, and b) meeting established technical criteria for selecting a high quality assessment?

## ***What we are used to in state assessment procurement***

There are two basic models for selecting an assessment contractor with which states, contractors, and state procurement specialists are quite familiar and comfortable. Prior to NCLB, the prevailing approach was to select an off-the-shelf norm-referenced test (NRT) from one of a handful of established vendors. Since NCLB, states have selected contractors to support the state in the development and administration of custom criterion-referenced tests (CRT). The level of state involvement in the development and administration process varies widely across states, but the fundamentals of the custom CRT procurement model are fairly consistent.

The most important difference between the Pre-NCLB off-the-shelf NRT and the NCLB era custom CRT models is the existence of the assessment at the time the contractor is selected. In the Pre-NCLB model, the test is fully developed and evidence of its technical quality is a significant component of the selection process. In the NCLB era model, there is no existing test when the contractor is selected (although in some cases there may be an item bank), therefore, evidence of the technical quality of the test is replaced by a demonstrated capacity to develop and deliver a custom assessment in the selection process.

---

<sup>1</sup> With thanks to my colleagues Brian Gong, for his input and thoughts prior to the production of an initial draft of this document, and Chris Domaleski, for his comments and feedback on early version of the document.

### *What is different with Smarter Balanced and PARCC*

As the 2014-2105 fiscal year approaches and states have to secure contractors and/or funds to administer their assessment programs, it is becoming clear that PARCC and Smarter Balanced do not fit nicely into either of our familiar procurement models. The fact that the multi-state, national consortium projects do not quite fit the routine models is not really too much of a shock. What is surprising, however, is that the complication is not due to the states trying to approach procurement as members of a consortium – although the consortium issue does provide some interesting angles of its own. The biggest complication is the point in the assessment development process at which the states are **first** seeking to select an assessment contractor for their new state assessments.

At this point in time, late fall 2013, we are in a sort of assessment limbo – in the Roman Catholic theological use of the word to denote being confined in a region on the border between heaven and hell, and not to suggest that we are happily engaged in a popular party dance. There is no question that at this time neither PARCC nor Smarter Balanced can be regarded as existing tests that would fit within the pre-NCLB procurement model. Specifically, it is not yet possible to evaluate the technical quality of either assessment. On the other hand, it appears that we are well beyond the point at which the states are selecting from among qualified contractors to support them in the development from scratch of a custom assessment; that is, the NCLB era scenario in which the technical quality of the assessment would not be a factor in the selection process.

How then should we view this procurement process as states select contractors and commit funds for their new state assessment programs?

### *Attitude is Everything*

As states within the Smarter Balanced and PARCC consortia begin to draft their assessment requests for proposals (RFP), we can borrow from the optimistic approach used on the Common Core State Standards (CCSS) website that classifies states as having “Adopted” or “Not Yet Adopted” the content standards (<http://www.corestandards.org/in-the-states>). For our purposes, we can distinguish between states that appear to have “Adopted” or “Not Yet Adopted” a specific assessment.

#### *Adopted*

States that have “adopted” the assessment will be able to indicate that clearly in their RFP. For those states, one can argue that the procurement process is consistent with the traditional NCLB era process of selecting a contractor to develop and administer a custom assessment, but that the contracting process simply has been divided into multiple parts. The argument in those states would be as follows:

- The state, in conjunction with other states, committed in spring 2010 to develop a custom assessment aligned to the CCSS.
- Since 2010, the states have issued a series of competitive RFP to select the most qualified contractor for each phase of the process of developing a custom assessment.
- The current administration RFP is simply the next link in the chain that will result in a high quality, custom assessment aligned to the CCSS.

## Contract Conundrum

### Not Yet Adopted

The argument is a bit trickier for those consortium states that we include in the category of having “Not Yet Adopted” a particular assessment. Note that these are not states that have left a consortium (e.g., Oklahoma or Georgia) or never joined a consortium (e.g., Texas, Virginia). These are states that have been participating fully in a consortium and *intend* to administer the consortium’s assessment, but for one reason or another cannot firmly commit to that assessment at this time. This includes the ever-growing list of states who must avoid using the term “common core” or naming one of the consortium-based assessments in their RFP because of political backlash against anything associated with the common core. It also includes states such as Massachusetts, whose Commissioner and Board of Education have made it clear that they have not yet adopted PARCC as their new state assessment and will not do so until its technical quality is established (<http://www.doe.mass.edu/boe/docs/2013-11/spec-item1-item2-motion.pdf>)

For states in this “Not Yet Adopted” category it is much more difficult to craft an argument that would

- support the sole source award of an assessment contract to administer either Smarter Balanced or PARCC; or
- lead to RFP evaluation criteria which do not include an evaluation of the technical quality of the proposed assessment (i.e., criteria which PARCC and Smarter Balanced are not yet prepared to meet)

The Massachusetts expanded field test or “*test drive*” approach is one promising solution to extending the custom development argument through the initial administration of the PARCC assessments. Large-scale pilot testing, or beta testing, of an assessment program has been a routine and critical component of the custom assessment process. Expending state funds for such activities prior to declaring the test operational could certainly be regarded as a continuation of the custom development process. Of course, that approach raises other issues outside of the procurement process such as funding and operating two assessments during the extended transition, maintaining accountability systems, etc., but those are issues for a separate discussion. Also, the “*test drive*” approach places much more emphasis on the technical and operational status of the new assessment than it does on an “adoption” decision. In other words, the implication is that we have adopted this custom assessment and will begin to use it when it is ready rather than the view that an adoption decision is deferred until after the assessment is developed. An analogy might be accepting delivery of a custom-built house from a general contractor – you will not accept the house and move in until all of your specifications have been met, but there is no question that, barring a catastrophe, you will move into that custom house that you are building at some point in the future.

A tried-and-true, but much more pedestrian, less intellectually appealing, and more ethically shaky approach is to craft an RFP with scope of work language, evaluation criteria, and response parameters that effectively eliminate all but the desired assessment and/or contractor from the RFP process. Although an effective workaround in the short term, when attempted without a supporting rationale this approach can appear to be simply a ploy to select a preferred contractor (or assessment) rather than as an attempt to build a custom assessment. Also, building such tight constraints into the RFP

process (or more so into regulations or legislation) can often cause unintended complications when the state decides to make changes to the design of the assessment.

### Not Yet Decided

In the ‘Not Yet Adopted’ category, we included states functioning within one of the assessment consortia and fully intending to administer the consortium assessment. There are other states for which the term “Not Yet Decided” is the best description of their current status. These states (e.g., Florida, Georgia) are considering an open RFP process that might result in the state developing a custom assessment, selecting a commercially available assessment, or perhaps even selecting PARCC or Smarter Balanced as their state assessment. Although such an open approach has some intuitive appeal, it also brings us back to the heart of our original dilemma: how to develop an RFP and evaluation criteria that allow for a fair and direct comparison of existing tests v. partially completed tests v. the potential and promise of a custom-developed test.

Although states may truly have “Not Yet Decided” whether to develop a custom assessment or select an existing assessment, it is likely that they have made some decisions that will inform the RFP process. As a starting point, a state might do well to fall back on the traditional approach of identifying features that must be included in their assessment program (e.g., alignment to the CCSS, assessment of deep levels of cognitive complexity, constructed-response items, computer-based delivery, quick turnaround time for reporting, rigorous achievement standards), features that they would like to have (e.g., adaptive testing, teacher involvement in item review or scoring, formative/interim tools and supports, a vertical scale) and constraints (e.g., cost, testing time). This will establish parameters for contractors to propose a method(s) that meets the state’s needs and for the state to establish criteria to evaluate contractors’ proposals. In fact, a state may enter the process of developing the RFP as a true “Not Yet Decided” state, and as a result of the process it can become clear before the RFP is issued that there is only one contractor and assessment that can meet the state’s needs. In that event, the state has may have moved into the “Not Yet Adopted” category.

Note that identifying “must-have” components of an assessment system is likely to rule out certain contractors or existing assessments, but that is fundamentally different than writing RFP specifications and evaluation criteria for the sole purpose of excluding all but one contractor or existing assessment.

### *Why is this consortium approach different from every other consortium approach?*

This is not the first time that states have joined together to form an assessment consortium and select a common contractor. Three oft-cited examples are the New England Common Assessment Program (NECAP), the American Diploma Project (ADP) Algebra I and II End-of-Course Exams, and the World-Class Instructional Design and Assessment (WIDA) consortium. Those three consortia successfully adopted and implemented three very different approaches to procurement.

The NECAP states independently issued RFP that required a common response to all participating states and jointly selected a common contractor. The ADP states issued an RFP through a single lead state, Ohio, and included riders and amendments that allowed other participating states to purchase tests through the Ohio contract. In both cases, however, the goal was to develop a custom assessment for

## Contract Conundrum

the consortium and an assessment contractor was selected at the beginning of the custom development process through a competitive bid process. As described above, in many ways it is the timing of the RFP in the development process that makes the current situation a bit more complicated.

The WIDA example is quite different than the other two consortia. With WIDA, the ACCESS for ELLs assessment is included as benefit of membership in the WIDA consortium. A key to states being able to effectively “sole source” the assessment through membership in the consortium is that consortium membership is routinely combined with adoption of the WIDA English Language Development (ELD) standards as the state’s content standards for English language proficiency. Consequently, although there may be other English Language Proficiency exams available that the state could consider in a competitive bid process, it is fairly certain that no other assessment system could claim better alignment with the state’s adopted content standards.

A similar process was followed when Maine joined the NECAP consortium prior to the 2009 test administration. One of the first steps in that process was Maine’s adoption of the NECAP Grade Level Expectations as their state content standards. Alignment with a unique set of content standards is one justification for a sole source award of the assessment contract, which leads us to some final thoughts on a yet to be played out consequence of multiple states adopting the CCSS, that is, a *common* set of content standards.

### *Irony, thy name is Common Core*

The biggest single difference between the current situation and previous assessment procurements is the existence of the Common Core State Standards (CCSS). The very thing that brought the states together in the first place, common content standards, may be a monkey wrench in the gears of the procurement process.

In the NCLB era, virtually all states have been developing custom assessments aligned to unique state content standards. As discussed above in the WIDA example, the combination of a custom assessment and unique content standards provided a buffer to challenges to the current state assessment. Notwithstanding claims of better technical quality, lower cost, or more streamlined administration, it was highly unlikely that an external test could show better alignment to the state’s content standards than an assessment developed by the state specifically to measure achievement of those standards.

With the adoption of the CCSS, however, that buffer is removed. There are likely to be multiple assessments at each grade level and content area that have been developed specifically to measure achievement of the same content standards. To what extent will, and should, a state be obligated to compare its current assessment against other offerings when issuing or renewing a contract? Would it be feasible or prudent for a state to attempt to mix and match assessment contractors to find the *best assessment* for each grade level and content area – if for example, niche or boutique contractors emerge who specialize in assessments in elementary school reading or middle school mathematics that are aligned to the CCSS? How often must a state consider other options – are there explicit benefits to continuity that outweigh advantages that a change might bring? How do those benefits fit within the current procurement process?

## Contract Conundrum

At some point as we weave through the layers of questions on the impact of the common core, the availability of multiple aligned assessments, and mixing and matching assessments across grades and years, our *contract conundrum* begins to fit the definition of a conundrum as a difficult or puzzling question having only a conjectural answer. Understanding the intricacies of the new assessment landscape and the ways in which that landscape interacts with existing procurement regulations, however, is more than an academic exercise. We have found in the past that the state procurement process is usually flexible enough to accommodate assessment needs, but only when those needs are fully understood and can be easily communicated.