U.S. DEPARTMENT OF EDUCATION

ASSESSMENT TECHNOLOGY STANDARDS RFI

Presented by:
Wireless Generation
55 Washington St.
Suite 900
Brooklyn, NY 11201
bids@wgen.net

January 17, 2011

Table of Contents

1.0 Cover Letter 3

2.0 Interoperability Recommendations 4

2.1 Interoperability to Support Innovation 5

2.1.1 Building and managing assessments 5

2.1.2 Working with and expanding on assessment results 5

3.0 An Open Interstate Assessment Bank 9

3.1.1 Functionality 9

3.1.2 Rationale 10

3.1.3 Leveraging existing projects 10

4.0 About Wireless Generation 11

1.0 Cover Letter

January 7, 2010

U.S. Department of Education
Steve Midgley, Office of Educational Technology
RTTA-RFI@ed.gov

Re: RTTT Assessment RFI Response

We are pleased to submit this response to the USED's Assessment Standards RFI.

We focus our recommendations on three key areas:

- The context for interoperability – what system boundaries make sense and why;
- Interoperability to support innovation – which domains are most likely to drive innovation; and
- Item, task and assessment management – recommendations for a shared platform to support this work.

Should USED desire any additional follow-up information, we would be glad to provide it.

Sincerely yours,

Josh Reibel
President & COO

2.0 Interoperability Recommendations

We support the overall goal of the RTTT Assessment Program – to create a new system of more effective and instructionally useful assessments. We also believe that maximizing interoperability of the assessments across technology platforms will drive innovation in the tools and services that support K-12 teaching and learning. We have a few strategic recommendations on interoperability, described below.

However, there is a key question that does not appear to be addressed in the RFI, nor fully established in either of the Assessment Consortia's proposals: what are the system boundaries? Which is to say, what systems, exactly, are interoperating?

One could imagine two extreme scenarios, with a hybrid solution being the most likely path to success:

- In a "thousand-flowers-bloom" scenario, interoperability is point-to-point. In the RTTT-A context, one could imagine 5 different state scoring and administration systems, each of which would have to be capable of feeding results to 3 different data analysis and reporting systems, each of which would in turn be expected to output data to dozens of state, local, and vendor systems to support instructional planning and human capital management.

- In a "monolithic" scenario, as much functionality as possible would be developed in a single system, either managed by or contracted by the Consortia. This system would have methods of delivering data out as required, but would aim to be a largely complete "stack" of services from the act of assessment development through to instructional delivery.

- In a "hybrid" scenario, one or two key components of the overall architecture would be established as a "platform," with which a variety of systems would be able to interact to access and deliver data and functionality.

We believe the hybrid scenario is most likely to be successful. The record of "thousand-flowers-bloom" interoperability is dismal across industries. With the exception of financial services (where significant resources are available and incentives are tied to transactions), we are not aware of a field in which interoperability standards have worked. Conversely, rolling as much functionality as possible into a monolithic solution is likely to constrain innovation.

We believe that a successful program will clearly identify and then architecturally define the intended areas of competition and innovation, stabilize those spaces by providing good APIs and strong data standards, and ensure that competitors and innovators in each area have timely and complete access to data to carry out their work. Doing so will greatly increase the leverage of the public, private, and philanthropic resources that will be invested to solve the next-next great challenge – how to use the data generated by the Consortia as both an input to and a measure of truly personalized learning.

2.1 Interoperability to Support Innovation

We believe that the interoperability standards work of the Assessment Consortia has the potential to drive innovation in two key areas: "before testing" (building and managing assessments) and "after testing" (working with and expanding on assessment results).

2.1.1 Building and managing assessments

We see two key benefits to interoperability in this space:

- Reduce switching costs for states – if all Consortia assessments are stored and accessible in a standard format, it becomes easier for states to take their assessment administration programs out to bid. States will then be in a position to determine the key areas for competition – reliability of online administration, for example, or quality of accessibility features.
- Reduce startup costs for vendors – if all Consortia assessments are stored and accessible in a standard format, vendors will be able to easily load the items or tests for a specific state's high-stakes assessment. These reduced costs can be expected to result directly in reduced pricing, particularly in an open competition.

To achieve these benefits, we recommend that USED and the Consortia:

- Make the "low switching costs" principle a key driver in the program plan. To accomplish this, each Consortium should consider requiring the final year of piloting and implementation in the development grant to be conducted using at least two administration and scoring vendors. The addition of a second administration and scoring vendor will be the "proof point" that the interoperability standards have achieved their objective.

In terms of the current standards landscape, we have experience with the QTI standard. While it addresses some of the key features required for item and assessment management, we found it to require significant extension. We believe that the objectives described above can only be achieved if those extensions (related to the management of assessments, linking of items to passages, etc) are consolidated into the revised specification.

We also have exposure but no experience with the APIP standard. Our preliminary review leads us to believe that it does an excellent job of extending QTI to address accessibility and deliver multiple-modality assessments. We believe this work should be incorporated or extended to accomplish the assessment- and item-management standards.

2.1.2 Working with and expanding on assessment results

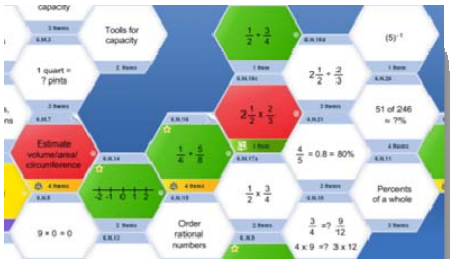
We believe the most important innovation to spring from the Consortia's work will be in the work done *after* assessments are completed. By ensuring that all student-level results data are readily available in a standard format and via standard methods, the interoperability standards and overall architecture will

drive competition in the **supplemental/diagnostic assessment, instructional planning, and human capital management spaces.**

Today, one of the greatest challenges to innovative technology developers in the K-12 space is the difficulty capturing and integrating the diverse data streams that are generated about a student. The best instructional and human capital management solutions need to know – in real time – how students are doing (assessment); what they are doing (instructional planning and delivery); and what adjustments can or should be made (coaching and professional development).

Because both Assessment Consortia are contemplating assessments over the course of the year, they will be generating data that will be vital to the delivery of truly personalized instruction. Online intervention providers will need to use the data to refine areas of focus. Diagnostic and progress monitoring providers will need to use the data to determine the best area of focus for each child and the baseline from which to look for growth or decline. Professional development providers – both in-person and online – will need to use the data to identify relevant supports that they can provide to individual teachers or groups of teachers.

Put more simply: advances in technology and instructional planning are pointing us to a future where student learning is personalized according to student need. The intra-year assessments collected as part of the Assessment Consortia's programs will be a crucial input for identifying those needs and adjusting instruction accordingly.



Our system for personalized learning, the mCLASS Platform, is a "So What, Now What" information system – a system designed to guide educators and administrators to a thorough understanding of what their data means (So What?) and help them develop and execute sound action plans that are grounded in the data (Now What?).

To impact student achievement, the system must cause something to change at the instructional core: the moment between a teacher, a student, and a task. That could be an administrator arranging for the most effective teachers to be in front of the neediest students; or a teacher identifying students who need scaffolding to read at the necessary level of text complexity. Teachers need direct access to 360 degrees of data for the latter.

That's why the mCLASS Platform includes honeycomb cells that are richly informational and truly actionable. For each standard, a teacher can fully understand a student's performance at the item level and in context of the assessment, the class, the district, and the state. Then, he or she can make a detailed action plan for full-class, small-group, or individual instruction to ensure each child is on the path to career and college readiness.

The mCLASS Platform (like its competitors) is a natural consumer of data generated by the Assessment Consortia.

To ensure a competitive and innovative environment in this space, we believe that the Consortia should:

- Define "student outcomes" standards with the same degree of definition and rigor of expectations as the item and assessment standards. These standards should cover, by item or task and by assessment:
    - Standard or standards addressed
    - Student performance (e.g. correct/incorrect; 3 out of 5; rubric score)
    - Student outcome (e.g. proficient, deficit)

  Especially at the assessment level, it is highly likely that a version of "business rule" definition will be required – e.g. different combinations of item scores or subtask performance may add up to similar "result" outputs. By exposing those rules and the underlying data, the downstream providers will be able to best triangulate on student ability.

  It is crucial that the performance and outcome data is tied to individual standards or sub-standards, even if the assessment did not collect information to make a definitive "outcome" decision on each standard. To put the data to diagnostic use, innovators will want to be able to combine the standard-by-standard data with that collected from other sources and use the combination of scores to place and plan for each student.

- Ensure that the results of an assessment administration are available rapidly (e.g. on the order of hours or days, not weeks or months), either directly from the assessment scoring vendor or from an intermediary technology layer, via standard APIs with appropriate security and privacy controls. A Service Level Agreement (SLA) should be established on this functionality.

We are familiar with the SIF framework and standards. SIFA been working on data standardization and flow in education for a number of years, and we believe that their emerging work on the student portfolio could represent the best standards developed to date. However, our caution on SIF is that to-date it has required district-level Zone Information Servers (or ZISs). Whether hosted or sited in a district, the ZIS is not a solution for every district and seems particularly poorly-suited for the largest. A

specification and protocol that is web-services-based, on which SIFA is currently working, is far more likely to be successful at scale.

To support the Consortia in this work, we believe that the U.S. Department of Education should take a strong hand in defining these standards and requiring full compliance with them. There is significant risk of something along the lines of a "browser war" taking place, whereby the vendors who win the assessment administration and scoring and/or data reporting contracts extend, enhance, or customize the standards and APIs and in doing so make it difficult for state or local data systems or other vendors to access and use the data. Winning the scoring contract for a state or for the consortia should not unduly advantage a vendor in other areas of the K-12 market.

3.0 An Open Interstate Assessment Bank

One component of the value stack that we recommend implementing as a platform is item and assessment management, via an Open Interstate Assessment Bank (OIAB). This is a natural and obvious area of inter-state and inter-consortia collaboration; an opportunity for dramatic cost savings, efficiency, and improved quality; and a key point of leverage to ensure successful interoperability elsewhere in the stack.

3.1.1 Functionality

The OIAB would be designed to serve both the work of the Consortia (e.g. developing and managing high-stakes assessments) and broader assessment work (e.g. developing and delivering complimentary diagnostic or screening assessments). It would have three key areas of functionality:

- Item and task management. A fully-featured repository for all types of assessment items and performance tasks to be used as part of the Consortia-developed assessments, which could quickly grow to support much broader collaboration to support diagnostic, interim, and screening purposes. Items or tasks could be created, browsed, searched, extended, rated, and shared. Items could be tagged according to license (e.g. public domain, creative commons, or proprietary). Proprietary content could be protected (perhaps even with a fee-for-access business model?) as desired by the Consortia or the creator of the content.

- Assessment compilation. The assembly of items and tasks into discrete assessments that determine student proficiency on specific standards at a given level, or that are designed to find a student's level in a particular set of standards. As with items or tasks, assessments could be shared, managed, etc.

- Access management and security. Because the Consortia assessments will be high-stakes, significant security will be required at the item, task, and assessment levels. As appropriate, content may be released (along the lines of current state processes to "release items").

Inputs to the OIAB would include:

- the content developed by and for the Consortia;

- released high-stakes assessment tasks from state assessment programs;

- district-created (or purchased) assessment content;

- test items or performance tasks developed by individual teachers; and

- items or tasks included in instructional programs, as released by the publishers.

Outputs from the OIAB would include:

- bulk export of items, e.g. for use in another diagnostic assessment system;
- discretely created assessments, either for administration and scoring in the Consortia assessment program, or for use by a screening or diagnostic assessment vendor;
- cumulative research data on the validity and utility of specific items or tasks in the context of the Consortia.

3.1.2 Rationale

Why establish the OIAB as a platform, rather than trying to create a market for item management systems?

1) Items and tasks are opportunities for innovation; endlessly re-exporting them is not. We believe the Common Core standards and the work of the Assessment Consortia will trigger an explosion in interesting and innovative new approaches to assessment. For each new task or approach that is created, instead of copying and re-copying the content, far more value can be generated by improving administration and scoring or by enhancing the utility of the results to inform instruction.

2) The Consortia need to consolidate data at some layer or layers in order to be able to achieve interoperability. To enforce some of the requirements established earlier in this response (e.g standards linking at the task level), the OIAB provides a simple test – data can be loaded in if it is compliant with the requirements; otherwise it is rejected at the point of upload or input.

3) The difficulty of transferring items or assessments from system to system has plagued the testing field for decades. We are aware of a number of states that elect not to use other states' items simply because the expense of importing or re-entering the items is comparable to the expense of developing new items. By providing access to assessment items and tasks in a mode of "input once and transfer anywhere," the Consortia are likely to provide incredible value (access to high-quality assessment content) at low cost.

3.1.3 Leveraging existing projects

Working in partnership with the New Schools Venture Fund and Achievement First, we have developed an Open Item Bank that implements much (but not all) of the functionality described here, though with a primary audience of district systems. We would consider releasing some or all of the software under an open source license, and believe our partners might be willing to do the same.

4.0 About Wireless Generation

We are a provider of innovative education data and assessment solutions. Our award-winning mCLASS® solutions are used in 50 states and overseas to deliver assessments, reporting and support. We currently serve more than 200,000 educators and 3 million students in thousands of school districts. Through our professional services, we enable educators and education leaders to make informed decisions based on the outcomes of the assessment and apply the appropriate methodologies to improve results through aligned, easily-delivered activities.