

Considerations for Next-Generation Assessments:

A ROADMAP TO 2014

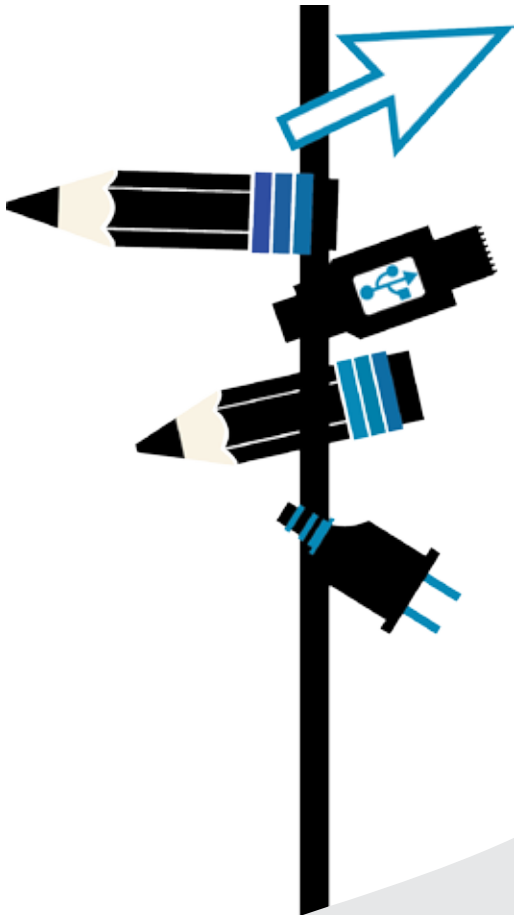




Table of Contents

Moving to Digital, Next-Generation Assessments	3
About This Roadmap	4
Why Go Online?	5
Five Step Roadmap	9
Conduct a Needs Analysis	10
Develop a Realistic Transition Strategy and Plan	16
Ensure Interoperability	18
Communicate Proactively	21
Plan for Ongoing Change	24
Call to Action	26
Case in Point: How Virginia Made the Leap	28
Case in Point: How Mississippi Built Success and Stability	32
Case in Point: Bridging Technology and Assessment in North Carolina	36

“The rate of change in the global knowledge economy is impacting everything we do—including everything we do in education. It is our responsibility as leaders to promote, identify and raise up this innovation in education and use it to change the landscape of teaching and learning.”

Gene Wilhoit, Executive Director of the CCSSO

1 Hines-Ward, Gretchen. (2010). Chief Justice John D. Minton Jr. and Gene Wilhoit to speak at Centre. Available at http://www.centre.edu/news/2010/march_convos.html.

Special thanks to the Council of Chief State School Officers (CCSSO), Digital Learning Now!, the Alliance for Excellent Education, the State Educational Technology Directors Association (SETDA), the Software & Information Industry Association (SIIA), Consortium for School Networking (CoSN), the Virginia Department of Education, the Mississippi Department of Education and the North Carolina Department of Public Instruction for invaluable contributions to the creation of this publication.

MOVING TO DIGITAL, NEXT-GENERATION ASSESSMENTS

Today an unprecedented convergence of forces is transforming American education. Included in this array are calls for the increased use of digital technologies and a clear need for more direct collaboration within and among states to prepare students for increased global competition. As noted in their 2010 Digital Learning Now! report, former Governors Bob Wise and Jeb Bush suggest that a critical part of this transformation will rest on effectively deploying technology to accelerate student achievement by personalizing instruction and assessment.

As far back as 2002, a U.S. Department of Commerce survey² reported that 90 percent of children between the ages of 5 and 17 used computers at home, school, or both. And in October 2010, the International Association for K-12 Online Learning released data showing that supplemental or full-time online-learning opportunities are available to at least some students in 48 states, plus Washington, DC.³ These numbers portray a wide-spread presence of technology in schools, but they fail to highlight the many challenges remaining for states seeking to maximize the benefits of using technology in the classroom. Most significantly, it can be time-consuming, costly, and logistically difficult

.....
2 United States Department of Commerce. (2002). A nation online: How Americans are expanding their use of the Internet. Washington, DC.

3 Wicks, M. (2010). A national primer on K-12 online learning: Version 2. iNACOL. Available for download at: http://www.inacol.org/research/docs/iNCL_NationalPrimerV22010-web.pdf.

to develop the infrastructure needed to support large-scale use of technology in schools.

Simply put, as Governor Bob Wise has stated, *“In just about every other facet of society, at work and at home, technology has transformed the way Americans go about their lives. Yet schools have been slow to embrace the transformative power of technology. Although computers are pervasive in schools, they tend to be used more like electronic textbooks—high-tech tools in a nineteenth-century system.”*⁴

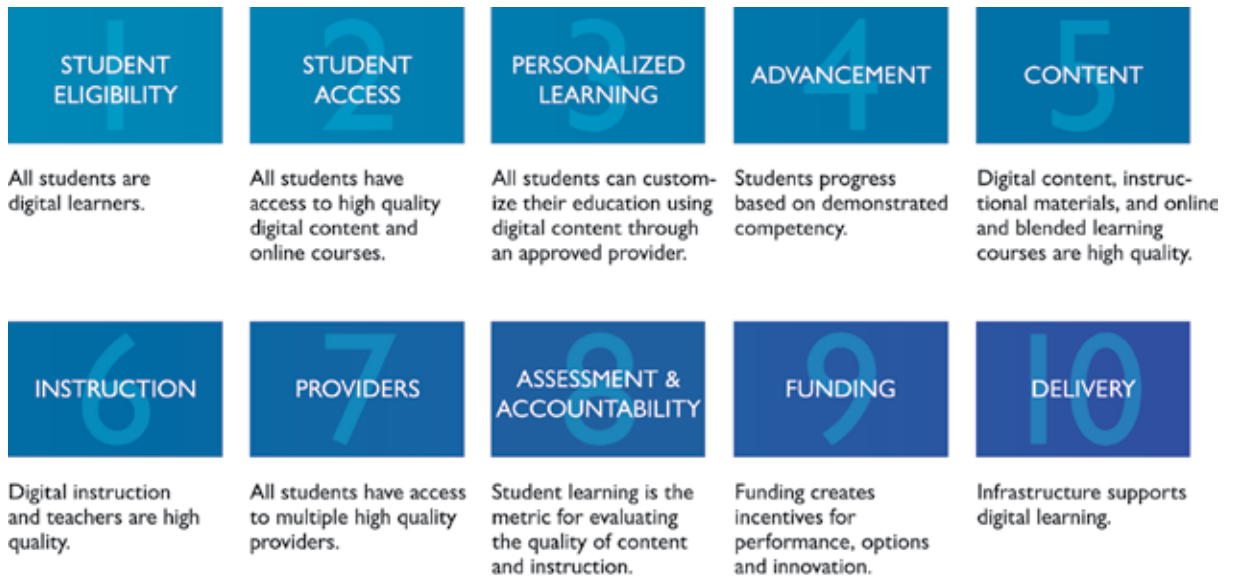
Producing graduates who will succeed in the current technological landscape can be simultaneously a top priority and a hurdle to overcome. Digital learning—given the right planning and support—can provide students and their teachers the utmost quality of instruction and, ultimately, level the proverbial playing field for graduates as they enter the workforce.

To help address gaps in funding and other resources in student assessment—through its 2010 Race to the Top Competition—the U.S. Department of Education (USED) awarded two state consortia grants to create new, comprehensive assessment systems: The Partnership for Assessment of Readiness for College and Careers (PARCC)⁵ and the

.....
4 <http://www.all4ed.org/files/OnlineLearning.pdf>

5 <http://www.achieve.org/parcc>

10 Elements of High Quality Digital Learning



* Source: Digital Learning Now!, December, 2010, Foundation for Excellence in Education

SMARTER Balanced Assessment Consortium (SBAC)⁶. The SBAC and PARCC consortia are now focused on establishing infrastructure and content for common, online assessments. They are building on the foundation of the Common Core State Standards (CCSS)⁷ that have been developed and adopted by more than 40 states. Central to the efforts of both consortia is the planned online delivery of assessments in 2014–15.

Independent of the consortia, most states have been involved in some type of online assessment effort, ranging from early stage pilot initiatives to full-scale, operational online assessment systems. In fact, at the Maryland Assessment Conference in October 2010 it was noted that 44 U.S. states have current computer-based testing (CBT) initiatives underway.

But the transition from today’s paper-based assessments to greater utilization of technology-based systems—in particular, the more innovative assessment methods and content needed to respond to the growing emphasis on college and career readiness—will be complex. States will invariably face challenges and uncertainties along the way. Careful planning and management at both the state and local levels will be critical to success.

⁶ <http://www.k12.wa.us/smarter>

⁷ <http://www.corestandards.org>

ABOUT THIS ROADMAP

This road map is intended to help states navigate—and mitigate—the many interdependent issues and challenges they will face as they transition to online assessments. Based on insight from states that have already made the transition and from assessment experts, this guide contains valuable lessons learned as well as five key steps to success:

- (1) Conduct a needs analysis
- (2) Develop a realistic transition strategy and plan
- (3) Ensure interoperability
- (4) Communicate proactively
- (5) Anticipate ongoing change

To provide the most accurate and relevant information about the state of online assessment, the roadmap was developed by a team from Pearson with support from education technology-focused associations and state assessment and technology leaders. Additionally, a public wiki website was created in the spring of 2011 to allow public comment and contribution from all interested parties.

“What you’re seeing across the states and in the education industry as a whole is broad recognition that there is so much more we can do for students to be successful, if only we better leverage the assessment-system capabilities that technology presents us.”

Jon Twing, Chief Measurement Officer for Pearson

WHY GO ONLINE?

In the past, testing programs have been criticized at times for their distance from classroom instruction and learning. Teachers feel testing is something done to them and do not feel current state of the art assessment reflects meaningful learning in the classroom. Furthermore, assessment has been accused of ‘narrowing the curriculum’ and watering down rich and complex student learning and teacher instruction.

Regardless of the truths to these claims, technology will enable a much fuller integration of instruction with assessment; perhaps even enabling the creation of individualized learning paradigms. More importantly, the use of technology will allow for the teaching, learning, and assessment of very rich and complicated problem solving activities, critical analyses, the generation of evidenced-based arguments, and the defense of such arguments—all skills required for success in college and the workplace.

Richer and more innovative item types.

Online testing, with its ability to employ technology for animations, simulations, and other advanced assessments of deeper understanding and complex reasoning skills, offers the ability to measure a fuller range of cognitive complexity. Tests can be designed with audio and video streaming that not only facilitates contextual understanding, but can increase students’ level of engagement. Audio and video streaming,

complex scenarios, and access to multiple data sources also allow for linking several items together to assess multiple aspects of a construct.

Existing capabilities such as click-and-drag allow students to respond to concepts more interactively. For example, students can complete a chart by clicking and dragging icons from a menu of response options to a chart location. These features allow for a deeper assessment of students’ understanding, by incorporating grouping, ordering, and other cognitive tasks relevant to assessing thinking skills and the depth of understanding of students’ knowledge and skills.

Directly tied to these richer, more innovative item types will be the need to upgrade the bandwidth capabilities of local technology infrastructures. According to a recent report⁸ commissioned by the Federal Communications Commission (FCC) to collect data on the current state of broadband connectivity in U.S. schools today, only 22 percent of district respondents said that their connection speeds completely meet their current needs.⁹ This is a critical point, because the types of items and assessments being designed for 2014 will likely

8 http://www.fcc.gov/Daily_Releases/Daily_Business/2011/db0106/DA-10-2414A1.pdf

9 http://www.fcc.gov/Daily_Releases/Daily_Business/2011/db0106/DA-10-2414A1.pdf



Technology-enhanced testing allows more innovative methods of assessing student performance and cognitive understanding. In this example, students perform an experiment while summarizing the results in their own words.

be much more media-rich and data-intensive. How much additional bandwidth they will require will depend on the number and nature of these innovative item types, but it is very likely that they will need more bandwidth than most current online assessments require.

Another key consideration when planning for dynamic items is timing and commitment to the online testing mode: In order to provide the right equity for the entire testing population, use of these technology-based innovative items may require full participation in online testing. When only some of the student population is able to test online, and make use of these items that cannot be delivered on paper, an obvious challenge to score comparability is presented. In essence, schools must first be able to fully transition away from paper testing, after which point the state can begin to fully realize the richer opportunity for the types of assessment items that only technology-based tests can deliver.

More efficient scoring capabilities. The robust, secure data management and delivery systems required by online testing allows faster turnaround of student score reports

and assessment data. Student assessment data are delivered more quickly, can be produced and communicated digitally, and can be used formatively to help plan or adjust lessons to better engage students and meet their needs.

Lan Neugent, Assistant Superintendent of the Division of Technology, Career & Adult Education for the Virginia Department of Education, highlights the data capabilities intrinsic to online testing, stating that the ability to quickly receive data on individual students regarding their understanding of key curriculum knowledge and skills is “*absolutely of paramount importance.*”

Better security. Less human interaction with materials results in greater security. Districts that conduct statewide paper-and-pencil tests must account for all answer documents and used test booklets, which contain secure content, as well as various other documents and pieces of paper used in administering tests. And before district staff can return materials to the scoring partner, they must box up everything, class by class, and then drive the contents to a central location, where they will be organized, resorted, re-boxed, and shipped.

“The district test coordinators love it. It’s less paperwork. They don’t have to worry about inventorying the test materials or losing things, so they like it.”

Jan Kirkland Hogue, Contract and Assessment Analyst,
Mississippi Department of Education

It’s not unusual for large districts to have hundreds of hired staff involved in this endeavor—resulting in a significant expense and the potential for security breaches or human error.

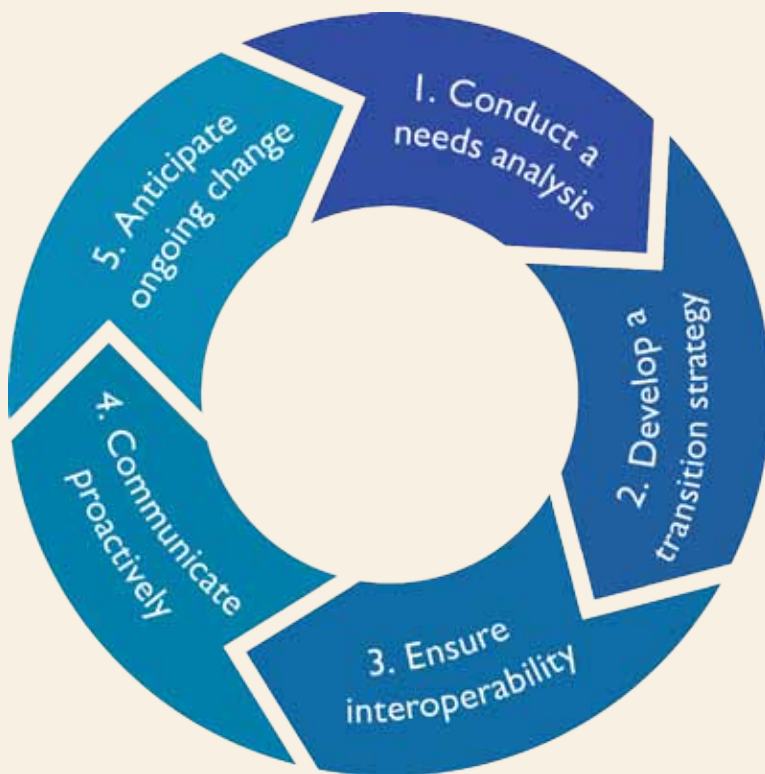
Greater equity. To date, accommodation strategies to support special student populations have often been considered after-the-fact, as retrofits. As such, they can compromise the validity of test results while still not best serving student needs. Better approaches, such as those based on universal design principles, can help design assessments that are inherently flexible and test students—even English learners and students with disabilities—on intended knowledge and skills.

For example, if an English learner does not recognize a word used to contextualize a math problem, one common linguistic accommodation is to allow that student to use a translation dictionary. But students are not all equally comfortable or adept using this support—especially if help is needed several times during the test—and they may instead guess at words. With a properly designed online assessment, we can allow students to simply click on words to read or hear the accompanying definitions, possibly even translated into another language. In this way, students have access to personalized, and private, supports, and test results gain validity.

Improved efficiency. Online testing uses electronic resources to eliminate or reduce the burden, labor, and waste associated with paper-and-pencil assessments, thereby lessening both labor effort and environmental costs. States with a stake in “green” and waste-reduction initiatives stand to see both reductions in paper needs—for test books, answer documents, ancillary materials, labels, and other materials—as well as in emissions and fuel required for transport of printed materials. This improved efficiency also has the potential to reduce overall delivery costs, although one important factor to consider during the transition is that delivery costs can sometimes increase temporarily, such as when both online and paper-and-pencil tests are delivered throughout the state.

Increased student engagement. Online testing provides wide opportunities for interactive experiences for students. Also, because students’ experiences are increasingly with digital media, online testing creates a more authentic, familiar, and engaging experience for students than paper-and-pencil tests. For example, students can manipulate and explore test questions, lending a more authentic experience and allowing results to reflect a deeper understanding of student knowledge and skills.

FIVE STEP ROADMAP FOR TRANSITIONING TO ONLINE ASSESSMENTS



Given the many research and technology-related issues, states must plan carefully and thoughtfully as they transition to online assessments. While specific actions within each state will vary, five main steps remain constant and are critical in helping to ensure proper planning and, ultimately, a successful transition to online testing:

- (1) Conduct a needs analysis
- (2) Develop a realistic transition strategy and plan
- (3) Ensure interoperability
- (4) Communicate proactively
- (5) Anticipate ongoing change

Each step includes a number of questions and considerations that are critical for states to work through as they make the transition to a fully functional online assessment system. The steps, along with related questions and considerations that are critical to states' planning efforts, are outlined in more detail in the following pages.



CONDUCT A NEEDS ANALYSIS



Before forging ahead with online testing, state education leaders must survey their needs and readiness—both figuratively, by creating a detailed assessment-design roadmap, and literally, by conducting a technology infrastructure survey of schools and districts. Doing so will guide states’ actions and focus their transition, while also helping policymakers determine the amount of resources and time required to get districts prepared to implement online assessments.

Create a detailed assessment-design roadmap. Looking ahead to a statewide online testing implementation can be daunting. States can position themselves for success by considering preferred content and item types, needs for retest opportunities, accommodations, comparability issues, legal defensibility needs, test security, and scheduling.

Content and item types. Starting with assessment content is critical to determining preferred question or item format. Different components of the content standards and curriculum will lend themselves to different assessment question formats. For example, some states might opt for gaming-style simulation items to teach and measure problem solving. Such a strategy could include computer-adaptive assessment components. Regardless, such a decision will drive considerations of the structure of the assessment program differently

than one that uses traditional multiple-choice type items. Such choices will also result in different infrastructure and assessment development needs.

One challenge related to moving content to online assessments concerns fidelity of assessment. Moving ahead of instruction by measuring students differently from how they were taught it is easy to do, but it is also a dangerous pitfall.

“It’s going to be hard to use, say, a simulated science experiment to measure a science lesson when in the classroom, the only exposure students had to the experiment was reading about it in a textbook,” Twing says.

The contrary is also true however. It has been argued by some that currently the only time students sit down to write a five paragraph narrative essay using a paper and pencil is when they engage a statewide writing test.

Also important when designing test content is avoiding assumptions that can impede accurate, authentic, and fair testing. The Universal Design for Computer-Based Testing Guidelines¹⁰ addresses item design vis-à-vis students’ varying perceptual, linguistic, cognitive, motor, executive, and affective abilities that aren’t intended for

.....
¹⁰ <http://www.pearsonassessments.com/udcbt>

measurement. As such, the guidelines support the development of tests that provide students better opportunity to demonstrate their construct-relevant knowledge and skills.

It's important to strive for a balance when designing assessments, so that tests leverage technologies in a way that decreases existing barriers without introducing new barriers.

Here are some other areas of assessment design that states should consider:

Comparability. One of the most critical uses of assessment results is in the ability to compare scores over time and across populations of students—within schools, across districts, and even, with the rise of common core assessments, from state to state. To support that, it is critical to ascertain that student performance results are not influenced by factors unrelated to their understanding of the content being assessed, such as might happen when some population of students take the test in a mode or environment more advantageous to their performance than the mode other students use for testing.

Policy makers should carefully plan out how much ongoing investment may be needed for critical research questions, such as the following: Will the presentation mode change over time for the same test (such as from the introduction of new technologies), making year-over-year

When creating an assessment design plan, states should explore in detail the following questions and considerations:

- How will the content I have selected help me accomplish my learning goals for all student populations?
- How will it help me improve instruction?
- How will it help me improve teacher efficiency, teacher efficacy, and teacher professional development?
- Which item types will provide the most useful accommodations?
- What comparability issues need to be considered and addressed upfront (e.g., mode, scoring, etc.), and how often will comparability studies need to be undertaken?
- What is the ideal length of the testing window?
- How will schools and districts manage the logistics around scheduling online assessments, likely with limited ratio of student to computers?
- Will summative assessment be given as a single exam or as sections offered at different times (e.g., by giving essays and other human-scored or open-ended sections earlier, or via through-course assessment structures)?
- What is the re-test policy and schedule, and how will these be handled?
- Will the assessment have time limits or guidelines (allowing schools to schedule multiple testing sessions within a single school day), and will additional time be considered as an allowable accommodation for students with unique needs?

“ [For our computer skills class], policy was established to open up the testing window to be basically all semester long, both fall and spring semesters ... it made sense that students shouldn't have to wait until the end of a semester or a year to take the assessment. ”

Randy Craven, IT Director, North Carolina State-TOPS

comparability suspect? Can performance-scoring decisions derived from paper administration be used for computer-administered tests? Does an online version need to look exactly like a paper test, or does some amount of difference (especially for targeting interfaces to best suit the specific mode) actually enhance true comparability? And as new computers and other devices, such as tablets, are introduced into school environments, are there any special considerations that may need to be accounted for in order to ensure the comparability of results by students on newer devices with the results of students who take the same test on the more “traditional” desktop or laptop computers?

A need remains to develop reasonable, generally accepted methods for addressing comparability issues, if comparability is desired. These questions are particularly important for the types of validity and reliability evidence collected and used to defend interpretations of the results from the assessment.

Test windows and test security. What is the desired length of the testing window? This decision may affect how many versions of the test (both paper and online) for which you must maintain test security. A computer-adaptive test may alleviate some of this dependency on fixed forms assuming a large item pool from which to draw. Steps like security bar codes for checking materials in and out, packing lists, counting

boxes and documents, physically returning all paper-based materials simply go away when testing online.

“[For our computer skills class], Policy was established to open up the testing window to be basically all semester long, both fall and spring semesters. Load management was critical to the success of the project. With the cumulative or spiraled design of the curriculum for computer skills, it made sense that students shouldn't have to wait until the end of a semester or a year to take the assessment.

That's in stark contrast to what folks want to do for an end of year or an end of course sequence. They don't want to test before the end of the series, and don't want to test too many days after the end of the series. So ... that puts a big burden on trying to design something that can accommodate [both tests].”

Randy Craven, IT Director, North Carolina State-TOPS

Re-tests. How are re-tests to be handled? For example, it might be more efficient to offer an “on demand” online retest for students who initially fail an assessment. Furthermore, it might be best from a security point of view to offer a computer-adaptive version of this online retest. If, however, the initial assessment was provided in a paper-and-pencil format, what steps are required to help ensure that scores from the retest are fair and comparable with the paper-based assessment? What is our plan for explaining to stakeholders (teachers, parents, students) how

“ Simply put, without continued and direct investment in broadband and educational technologies, education reformers are asking schools to improve, innovate and compete with one hand tied behind their back.”

Doug Levin, Executive Director of the State Educational Technology Directors Association

we know scores are comparable? If an adaptive test is used, how will we be able to explain that a student who failed the adaptive retake after only 21 questions would not necessarily have passed had they taken the 60 items that were in the paper-based test? These issues can be addressed, but they are complex and need to be factored into the planning process.

Class scheduling. Schools will have a limited ratio of computers to students if the current summaries of school infrastructure are correct. Therefore, each school might need its own plan for scheduling computer time to accomplish testing within the test window. Or, more resources may be needed to support the schools in accomplishing the assessment task within the testing window.

Conduct a technology infrastructure survey of schools, district, and states. Assessing states' and schools' resources and readiness is critical for success in moving and maintaining online assessment or instruction. One way some states have gauged readiness is through a survey of school and state leaders. Such a survey is also a useful tool for media relations, legislative briefings, and general knowledge about how and why investments in technology or infrastructure are being used.

Additional resources for developing an assessment-design roadmap are available for download from the interactive version of this roadmap or are available directly from the Next Generation Assessments website at www.PearsonAssessments.com/NextGen.

- [Thoughts on an Assessment of Common Core Standards](#)
- [Universal Design for Computer-based Assessments](#)
- [Consideration for Developing Test Specifications for Common Core Assessments](#)
- [Thoughts on Linking and Comparing Assessments](#)
- [Considerations for Performance Scoring When Designing and Developing Next Generation Assessments](#)
- [Automated Scoring for the Assessment of Common Core Standards](#)
- [Designing and Operating a Common High School Assessment Program](#)
- [Some Considerations Related to the Use of Adaptive Testing for Common Core Assessments](#)
- [Transition Planning for Next Generation Assessments](#)

.....
11 Aronowitz, Scott. (2/24/11) The Journal. National Broadband Map Suggests U.S. Schools Need More and Better Investment in Technology Infrastructure. <http://thejournal.com/articles/2011/02/24/schools-need-more-and-better-investment-in-technology-infrastructure.aspx>.

When planning for technology infrastructure surveys, states should consider these strategies and questions:

- How will the content I have selected help me accomplish my learning goals for all student populations?
- Can existing survey instruments be leveraged? What are the best methods for soliciting important information from key groups?
- Is a district-level survey sufficient, or will critical information be needed from each individual school?
- Who are the key personnel within districts or schools to help answer survey questions (e.g., administrative, assessment, or technology staff)?
- How many computers within the district or at each school meet technology requirements for an online assessment and can also be made available for student use during assessments?
- Given a specified window for testing, are there enough computers and lab times for all students to complete testing?
- Is enough network bandwidth available to support the content and number of students testing simultaneously? Do schools already have proxy-based or other mechanisms to reduce the demands on bandwidth?
- How often do schools refresh or replace their existing computers and other hardware?
- Does the school have future technology-implementation plans that could materially impact any of the above questions (e.g., a school that may intend to replace computers of one particular operating system with mobile handheld computers using a different operating system)?
- What other infrastructure or logistical issues may need to be assessed at the local level (such as the need for electrical power upgrades or other location-specific challenges)?
- Does the survey need participation from all local entities (which may require force of legislative or policy mandates), or is a representative sampling of districts sufficient to provide information that can be extrapolated across the state?

There are also facility-related considerations that should be accounted for, such as having adequate electrical power in the building, and having appropriate cabling, outlets, and jacks in place. *“If you don’t have the computer wiring, all the bandwidth and everything that you need ... in place, you’re doomed to fail,”* says Virginia’s Lan Neugent. A common challenge for schools is verifying their connectivity to the Internet, level of technical assistance, or firewall protection. As such, a survey is also a good way to establish local protocols, procedures, and investments.

Conducting readiness assessments of district and school leaders can help states get a more complete picture of their technology infrastructure. Moreover, state assessment and technology leaders and policymakers can provide valuable insight about state policies, priorities, and budgets as states initiate the planning process.

These additional resources for developing survey instruments are available for download from the interactive version of this roadmap:

- [Sample school-level survey template for online readiness](#)
- [Sample district-level survey template for online readiness](#)
- [Sample state-level survey template for online readiness](#)
- [Texas Legislative Report – An Evaluation of Districts’ Readiness for Online Testing](#)
- [Transitioning to Online Assessment in North Carolina](#)
- [Transitioning to Online Assessment in North Carolina Addendum](#)





2. DEVELOP A REALISTIC TRANSITION STRATEGY AND PLAN

After conducting a thorough needs analysis, states can decide on appropriate strategies for moving to online testing. Each strategy should detail expected and preferred steps and percentages of online testing and include contingencies for external conditions, including power or network outages.

Create a multi-year transition plan. The most consistently successful strategies for transitioning to online testing have typically employed stepped or phased transitions, often over multiple years. This allows districts and schools within the state to gradually build the necessary infrastructure and also to grow confidence and comfort with technology-based assessments within their staff and parent populations. This is particularly useful in communicating to teachers, parents, and students that change is underway and will be conducted in a controlled manner. As schools become more comfortable with setting up and running their tests, they can more readily expand outward.

Some strategies for phasing in online testing include beginning with one grade or subject at a time, with special populations (such as English Language Learners), or with exit-level graduation exams.

For example, Virginia, which started the move to online testing in 2001, began online testing only in high schools with available computer labs. Nearly ten years later, the state has become highly

“One of the really good things we did is we started this as a high school initiative ... because a lot of the high schools had business labs and other resources we could use for the testing while an independent capability was being built.”

Lan Neugent, Assistant Superintendent of the Division of Technology, Career & Adult Education for the Virginia Department of Education

comfortable with online testing and is poised to design tests specifically for online use. Virginia is one of the first states to make an almost full transition to online testing, and now conducts virtually all of its high school testing online, as well as about half of its elementary school testing. In 2010, the state delivered over 2.1 million tests online.¹²

During both the planning and the execution of any transition, it is critical to bring together assessment, learning, and technology staff, at state, district, and local levels. Each of these groups brings to the table a critical knowledge base that must work in unison. In particular, collaboration from both groups may be crucial to identifying key policy or security considerations that may need to change in the face of technology-based assessments. For example, many districts employ technologies in their computer labs that allow a teacher's computer monitor to mirror what's on a student's screen – but these types of technologies would typically need to be expressly forbidden as part of securing the assessment environment during testing.

Other considerations for states to include in transition plans are setting reasonable goals around cost, effort, and expected changes. At the same time, many states may need to maintain both paper-based and online systems until full transition occurs.

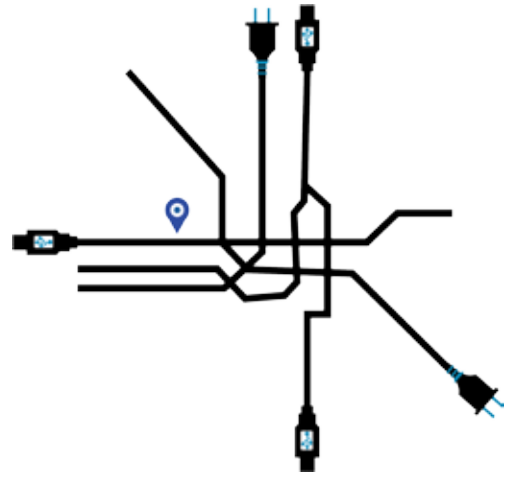
These additional resources for developing a transition strategy and plan are available for download from the interactive version of this roadmap:

- [Comparability of Computer-Adaptive and Paper-Pencil Tests](#)
- [Next-Generation Assessments website](#)
- [Online Testing Security & Policy Recommendations](#)
- [Case in Point: How Virginia Made the Leap \(at the end of this paper\)](#)

.....
¹² http://www.doe.virginia.gov/testing/online_testing/images/online_sol_test_chart2.jpg

When creating a multi-year transition plan, states should consider the following questions and considerations:

- Over what period of time will online assessments be phased in? What are the timelines for the technical implementation? What are the implementation strategies to ensure successful transition (phased in by grade, subject, low-volume tests, special populations, others)?
- Will the districts and schools be more successful by focusing on specific grades or schools (such as a transition that first moves high schools online, followed by middle schools, followed by elementary schools) or by focusing on specific subjects or other subpopulation groups (such as moving an ELL test online first, followed later by assessments for larger student populations)?
- Will both online and paper-based versions of the assessments be available during the transition process? If so, will districts be allowed to decide their own ratio of paper vs. online testing, or will that require state-level participation and approval?
- Are state and district assessment and technology directors equally committed to—and involved with—the transition planning to online testing? Is there a process and communication plan in place to support this joint effort?
- Contingencies for power outages, network outages, and other possibilities should also be addressed.



3. ENSURE INTEROPERABILITY

Interoperability is about portability. If testing vendors—development, delivery, and scoring professionals—adhere to the same technical standards, they can seamlessly share the components and resources that make up an online test. If student information is managed using clearly defined standards, student data can be utilized across all the integrated systems used by a state department of education or a school district. In fact, without standards of interoperability in place, states may be putting online test delivery in jeopardy. As a result, unnecessary money and time can be spent correcting incompatibility problems.

Jon Twing compares interoperability standards to the current standard of ‘plug and play’ in computer hardware: *“There can’t be pieces and parts that are disjointed. They’re going to have to be seamless or integrated, or at least able to exchange information freely. We see this with the evolution of personal computing from installation, maintenance and upgrading of devices drivers to the current standard of ‘plug and play’ where technology realizes the device and/or version and accommodates for its use. Assessment and instruction of the future must be ‘plug and play.’”*

At the 2010 symposium of the Software & Information Association, in collaboration with ASCD and CCSSO, 89 percent of attendees identified, as a next step toward a model of personalized learning, the “support [of] public-

private partnerships to advance key technologies, including common metadata and technical standards needed to enable the interoperability of various applications, data, and content resources to form a more seamless, integrated learning platform.”¹³

Ultimately, successful interoperability planning requires a thorough state-specific needs assessment coupled with the development of well-considered industry standards.

Evaluate and understand state-specific portability and interoperability needs.

The interoperability standards created and chosen by states will effectively formalize the capabilities and limitations to which all test development is subject. States should verify that the standards satisfy anticipated needs for all test question types and address data longevity.

To help ensure interoperability, states should assess the extent to which assessment items and even student and teacher performance data may need to move freely between different state, district, and third-party databases and

.....
¹³ Wolf, Mary Ann. “A Report From the 2010 Symposium.” *Innovate to Educate: System [Re]Design for Personalized Learning: A Report From the 2010 Symposium*. Washington, DC: Software & Information Industry Association, in collaboration with ASCD and the Council of Chief State School Officers, 2010. 7.

delivery systems. In the process of evaluating state interoperability needs, it is important to bear in mind some of the limitations inherent in choosing a defined set of standards:

- Standards may not be universal.
- Multiple standards may need to be employed. If a single set of standards is followed, and a state finds that more, or different, constraints need to be placed on test design, then a second set of standards may also be required.
- Standards need to be allowed to evolve over time. When a standard becomes insufficient because of new developments, the standard will likely need to co-evolve, be supplemented by another standard, or be replaced.
- Standards must be agreed upon by all stakeholders and, once chosen, will limit what the state can do and still have interoperable tests. Thus, the outcome will need to appeal to a lowest common denominator and may not feel as innovative as some stakeholders would like.

The key principle around interoperability standards involves mutual agreement on placing restrictions and constraints over content and technology. States, vendors and other parties who create and agree to adhere to these standards are implicitly agreeing to restrict their freedom to create or employ new and novel technologies, structures, or other mechanisms

States should consider the following when choosing interoperability standards:

- How will the content I have selected help me accomplish my learning goals for all student populations?
- How will it help me improve instruction?
- Will the state need to transfer item and other assessment-content data between multiple creation or delivery platforms, or between multiple assessment vendors?
- Will the state need to transfer assessment results or other student-teacher performance data between multiple data warehouses, portals, or other systems—either at the state level or to and from local, district, or higher-education data systems?
- Interoperability standards should minimize technology constraints and be free of any ties to particular hardware, systems, or technologies (such as Flash, HTML-5, etc.).
- Interoperability standards should allow for items and other content (such as images, animations, or question content) to be kept separate from formatting and other conditions that may need to change based on usage, policy, or technology (such as font type, size, position on screen, etc.).
- Interoperability standards need to support a wide range of content and be extremely flexible and accommodating of the widest possible types of images, audio, animations, video, and other content types.
- Interoperability standards should explicitly recognize and support a wide variety of assessment accommodations and structures (from tightly-controlled paper-based assessments, to dynamic and adaptive assessment systems), and need to allow for integration with other instructional content or data-management standards, such as MathML, SCORM, and SIF.



for the express purpose of gaining the value of portability between multiple systems. For this reason, it is critically important for the state to fully understand the standards—and the constraints—that are established, and to also understand the mechanism whereby the standards can grow and evolve in response to newer needs or technological capabilities.

Take an active role in creating interoperability specifications.

To avoid costly testing program problems, state leaders should work to establish interoperability standards that appropriately support current and future online testing capabilities. With a clear understanding of their own needs for interoperability, states should actively participate in industry-wide collaborations on interoperability standards. For example, the federally funded Accessible Portable Item Profile (APIP) project is a joint effort of stakeholders including a state consortium led by Minnesota and IMS Global, the creator and maintainer of the Question and Test Interoperability Interface (QTI) specification. Industry groups currently working to establish effective specifications in addition to IMS Global include the School Interoperability Frameworks (SIF) and the Post-Secondary Electronics Council (PESC). The inherent tension between the need for interoperability and the desire for innovation can only be resolved by active participation across the spectrum of stakeholders.

These additional resources related to interoperability are available for download from the interactive version of this roadmap:

- [Next-Generation Assessment Interoperability Standards](#)
- [Pearson's response to the USDE Request for Information on Assessment Technology](#)
- [IMS Global Question and Test Interoperability Specifications](#)
- [School Interoperability Frameworks Association](#)
- [Postsecondary Electronic Standards Council](#)
- [Accessible Portable Item Profile \(APIP\)](#)



4. COMMUNICATE PROACTIVELY



As with any successful technology implementation, communication and training efforts are critical. States should not just share information when asked, but must build and implement a communication and training plan that includes input from all stakeholders: administrators, teachers, IT staff, families, and the public. Collaboration across state lines can also be highly valuable, particularly allowing leaders to adapt lessons learned to inform the unique needs of their own state.

Likewise, it is critical that teachers be involved from the beginning. When teachers are highly informed and supported, the broader testing plan positively affects instruction and learning. To this end, future assessment systems will not only provide accountability data but will also provide access to information that will support and enhance instruction and student learning.

Create a forum for sharing technical knowledge at the state and local level.

One way to help ensure that all interested and involved voices are heard and that all perspectives are considered is to establish a statewide technical forum to facilitate inter-district discussion of assessment-technology needs, issues, and topics. These can exist as online message forums, in-person conferences, regional or statewide training events, or web or video-based conferencing opportunities.

“One lesson we learned is that you can’t communicate enough,” says Neugent. “We had schools in the mix. We had policymakers in the mix. We had internal people in assessment and technology instruction in the mix. And we tried to communicate with groups together. It wasn’t a case of, well, that’s their responsibility, or that’s that other office’s responsibility. You need to have people talking back and forth and being free to provide input into what goes on.”

Sarah Susbury, Director of the Office of Test Administration, Scoring & Reporting for the Virginia Department of Education, was involved with the Virginia transition from the start, and says communication not only alleviates existing fears but that it halts production of additional fears or rumors. By sharing information with principals, administrators, and central-office administrators, among others who were nervous about entering a high-stakes testing environment, Virginia leaders were able to answer questions proactively.

Further, throughout the transition process, milestones will occur: successes and mistakes will invariably be noted, providing excellent opportunities to share lessons learned. While some schools may wish to repress perceived errors, they will surely help others by relaying potential pitfalls, thereby increasing efficiencies and decreasing costs down the line. In this way,

States should address the following considerations as they assemble a communication and training plan:

- What is the preferred mechanism or forum within the state to best facilitate inter-district discussion of assessment-technology needs, issues, and topics (e.g., online discussion groups, in-person conferences, webinars, etc.).
- What are the preferred opportunities and mechanisms for leveraging existing state/district communication strategies and opportunities to keep teachers, administrators, parents, and policymakers well informed?
- Does the training plan sufficiently encourage joint collaboration between both assessment and technology staff members?



mistakes made in one district do not need to be repeated in another, and key lessons or successful best practices can help form a model that other schools can adopt.

Widespread communication has yet another benefit through indirect endorsement of online testing: Districts protect and support each other. When early adopters of the system discuss their success stories and the benefits of online testing, other schools listen. This helps states' end goal of full adoption. Similarly, when local schools fix problems or streamline procedures, other schools might benefit from such knowledge.

Finally, involve colleges and representatives from the workplace. By creating a feedback loop with these groups, Neugent argues, states will better be able to make adjustments to technology and their instructional paradigm.

“What we learned from talking with our districts about some of the keys [to successful online testing was that] the number one for most of them was the collaboration between testing personnel and technology personnel ... After that, I think it was communications. Not only between those groups, but [also] with the teachers in the school. [It was important to make] sure that the teachers were aware of what their options were, that they had the training that they needed to be able to administer the test online or paper/pencil, and [that] everyone in the school knows what time the testing is happening, where it’s taking place, and where the students are supposed to be, so you get people in the right places.”

Kayla Siler, Policy and Strategic Planning, North Carolina Department of Public Instruction

“ What we learned from talking with our districts about some of the keys [to successful online testing was that] the number one for most of them was the collaboration between testing personnel and technology personnel. ”

Kayla Siler, Policy and Strategic Planning, North Carolina Department of Public Instruction

Even with regularly scheduled communication forums and with everyone on the same page, great change can be accompanied by fear or trepidation. With paper-and-pencil testing serving as the staple in education for over a century, educators and education leaders are bound to resist, or at least be skeptical about, change. Some might even prefer to stick with the status quo even if they know change is required because change requires effort and entails risk.

Create a training plan for both assessment and technology staff within the state.

Besides involving all affected parties in the process and keeping them in the loop, a good way to raise comfort levels is to provide ample training opportunities. To do this, states should create a training plan that covers assessment, instruction, policy, and technical components. To further encourage strong partnerships at the local level, states should encourage or create opportunities, especially for joint assessment and technology staff, which will ultimately be integrated with instruction.

Examples of professional development for teachers include high-quality online courses; social networks for educators; live telecoaching sessions; electronic-capture opportunities for distance streaming of materials, plans, and instructional-practice vignettes over high-speed networks; and discussions on instructional

practice between candidates and instructors, candidates and mentors, and mentors and instructors.

Training should include opportunities for assessment and technology staff to learn together at the local level, side by side, with instructional staff. Says Jan Kirkland Hogue, *“Training for all new users, or users who have a little experience but are still kind of rusty—that is key.”*

Most importantly, testing and technology coordinators need training and support to help them lead other staff in transitioning. The more that affected contributors can come together, the more likely that cohesion will quickly follow.

These additional resources for creating communications and training plans are available for download from the interactive version of this roadmap:

- [Sample 96-hour Checklist from Virginia Transition](#)
- [Collaborative website for the Next Generation Assessments Roadmap](#)
- [Case in Point: How Mississippi Built Success and Stability \(at the end of this paper\)](#)
- [Case in Point: Bridging Technology & Assessment in North Carolina \(at the end of this paper\)](#)



5. PLAN FOR ONGOING CHANGE

Do not expect the transition to online testing to reach a final moment where nothing else needs to happen. The change to any technology-based assessment is not a single-point event, even when spread out over time. Instead, because technology is so wonderfully dynamic, states should expect that their next-generation assessment programs will also become equally dynamic to adjust to technology's ever-changing capabilities, products, and markets. This is a good thing because it means we are not slaves to technology but rather, our delivery platforms for instruction and assessment, once fully online, can be adapted and repurposed for the best efficacy and fidelity in both instruction and assessment.

Establish a plan for ongoing needs analysis.

To continue to support assessment goals appropriately, states should establish a plan and timeline to survey technology and infrastructure regularly. This plan should include periodic reexaminations of assessment needs, especially to reflect new capabilities and assessment methods. It should also include components to help gauge if the instructional and assessment online systems are being used in the manner that best meets the learning objective.

Without a set plan in place, states will likely have trouble keeping up with technological changes and will soon find that they are no longer current. The trick is to balance efforts between adopting what is new and supporting what has been established.

Consider the computer mouse, which did not accompany desktop computers when they first entered schools. Now, of course, a mouse with every computer is standard. Eventually, this too may change, and assessment programs will have to adapt accordingly. Also consider how the average size of computer displays in schools has fluctuated up and then down as monitor technology has changed, as flat-screens have replaced CRT screens, and as laptops and now netbooks have replaced more traditional desktop computers.

And now the world of mobile and handheld computing is also moving quickly into schools and classrooms. In the 2010 Federal Communications Commission (FCC) survey of schools, 60 percent of respondents indicated that they were either currently using handheld devices for academic or educational purposes or had plans to introduce them within the next three years. And the Consortium for School Networking (CoSN) recommends that districts find more ways to leverage student-owned devices as a way to increase access to technology when facing further cuts to technology budgets.

With their generally smaller screens and touch-based interfaces, handheld devices such as tablet computers could become one of the first disruptive elements to states' plans around online assessments, even as early as the 2014–15

school year. States who wish to utilize these new, portable, and relatively inexpensive devices will need to incorporate into their assessment plans contingencies to fund and develop systems for the new hardware.

Experts in educational technology recommend adopting systems that are open and based on industry standards as well as connecting with companies that work with open systems, or that are able to provide systems that will stay current with technology. The reverse is also true: Avoid partnering with a single hardware provider or one that relies on one particular technology to support learning systems, as this may prove to be a short-sighted solution.

Preparing for ongoing change is dependent on the ability to ensure continuous funding. Technology funding from states must remain a budget staple, for if it dissipates, online testing programs and the broader programs for online instructions to which they are tied may also collapse. Many states and districts establish a standard refresh policy dictating that computers will be replaced after a set number of years in service.

“It doesn’t take very much to become out of date very quickly,” Virginia’s Neugent says. “So you’ve got to do everything you can do to make sure that whatever funding you have is available for this and is continuous.”

These additional resources to help plan for ongoing change are available for download from the interactive version of this roadmap:

- [Texas Legislative Report – An Evaluation of Districts’ Readiness for Online Testing](#)
- [Transitioning to Online Assessment in North Carolina](#)
- [Transitioning to Online Assessment in North Carolina Addendum](#)
- [A Study on the Use of Netbooks in K12 Assessment](#)
- [National Education Technology Plan 2010 \(USDE\)](#)



When thinking about planning for ongoing change, states should plan around the following questions and considerations:

- How can future surveys best build on initial survey information to help the state stay abreast of emerging technology trends and other changes that may be happening or anticipated at the local level?
- What is the most appropriate level of effort required at the local level to best balance the state’s need for information against the cost and time impacts to district personnel?
- What are the realistic hardware options for states and districts to consider in the future for widespread use (looking one, three, and five years out)?
- What additional technology or infrastructure needs might districts wish to see in the online-assessment delivery engine in order to better meet their or their students’ needs regarding online assessments?
- Are state policymakers prepared to redirect funds toward or otherwise invest in technology education to support online assessments and the broader instructional-improvement programs to which they are aligned?

CALL TO ACTION

Work together: Teaming is everything!

Moving to an online testing system is a huge undertaking—one that can only be completed successfully with significant funding, local buy-in, and assistance from a strong and dedicated team.

States heading into online testing need to collaborate with experienced partners who have experience helping other states move to online testing and who can lend that knowledge and support locally, to help every school in the state be successful. Teaming with other states can also prove helpful, as many are still working through their own issues and value the perspectives of their peers.

Finally, national associations and consortia are essential to creating online systems.

Here is why:

- They promote longer-term stability.
- They do not ignore specific local needs.
- They do not ignore key needs for flexibility.
- They must monitor and maximize member investment.
- They understand the importance of having an integrated instructional and assessment system to ensure success.

“I would highly, highly recommend that any state that’s considering this either join a consortium or find a vendor that they can work with,” Neugent says.

Don’t be discouraged: Success breeds success.

The road to online testing can be a long and sometimes bumpy one. By focusing on and building from incremental successes, states and their districts will grow their programs with confidence. And instead of imposing the unrealistic goal of avoiding problems, states should instead include problem-resolution planning as part of their transition to online assessments.

As someone who’s been through the transition, Jan Kirkland Hogue says, *“You cannot imagine the different things that can come up with online testing. For instance, the very first time that we used it, we had a meteorite storm that completely shut down the whole system. Now, who would have ever thought that would have happened. It hadn’t happened in 400 years, and the very first day, that’s what happened: in fact we couldn’t even finish testing that day. So I would say expect the unexpected.”*

Help ensure consistency between instruction and testing. One way to make online testing successful is to help ensure consistency between instruction and testing. The modes by which students learn new materials should be consistent with the mode in which they are assessed. In other words, just as athletes must “practice the same way as they play” students must learn the same way they are assessed. This is something that, arguably, has been missing from assessment reform in recent years.

For example, if students use a specific brand of graphing calculator in the classroom, will they have access to a similar, if perhaps virtual, type of technology on the related test? It could be difficult for students to transition their skills and abilities from the functions and design of calculator that they use in the classroom to a new one that they use once during testing. The reality of this scenario affects the validity of how test results are interpreted.

Focus on the student. Above all, states should remain focused on the end users: students. Online assessments provide opportunities to better align and integrate curriculum, instruction, and assessment. And with high-quality technology-based classroom-assessment practices in place, states can more easily create and enhance a variety of assessments, including

- End-of-year, on-demand summative assessment;
- Frequent, on-demand, formative assessments; and
- Portfolio and other performance-based assessments.

“Make sure that, ultimately, the focus is still on the student and not the technology or the hardware, the servers, or the software that run behind it,” says KJ Singh, Chief Technology Officer for Pearson. *“To me, it’s all about if a teacher or a student doesn’t even have to think about what they are using—if the technology becomes second-nature. That is how success will be measured.”*



CASE IN POINT: HOW VIRGINIA MADE THE LEAP

Virginia began the transition to online testing in 2001, starting with high school students and one large-scale test. Today, the state is almost 100 percent online. Below is the story of how that state's program arrived at this point and of lessons learned along the way, as told by Lan Neugent, Assistant Superintendent of the Division of Technology, Career & Adult Education for the Virginia Department of Education; Shelley Loving-Ryder, Assistant Superintendent of the Division of Student Assessment & School Improvement for the VDOE; and Sarah Susbury, Director of the Office of Test Administration, Scoring & Reporting for the VDOE.

Shelley Loving-Ryder: One very important aspect in Virginia was that there was a commitment to move to online testing from all parties, from the governor and the state superintendent, to the legislature. When you are moving to something that is this new, it is important that everyone be committed.

Sarah Susbury: Buy-in from the community was significant, too. Students go home every day, and their parents hear about the changes. And school board members are asked about them. So we put in a lot of legwork in terms of communication and hand holding. And we made sure we had early successes so other schools felt good about the change and were motivated to put energy and effort into going forward themselves.

Getting down to business

Lan Neugent: The first thing we did was to develop architectural guidelines. And that became critical down the line because one of the things we discovered about the technological point was that we had to have some uniformity of technology throughout the state.

As we were going through, we discovered that we were not allowing school divisions to have direct access to the test. [That] really caused bandwidth issues, and so we put in place a proctor-caching scenario where schools could basically download tests, have them just like they would a paper test, on a server or on a lab server, or something of that sort. And then we developed a whole protocol for them to get rid of the test, so that they were not on machines in the future. Plus, Pearson encrypted all of those tests for security. To date, I don't think we've ever had any kind of a breach.

After establishing the architectural guidelines, we asked schools to certify that they had met them by ensuring that the computers they were going to be using, and the bandwidth, and everything else was all in place. After they self-certified, we required them either to certify with some kind of software that proved what they were saying to us was true, or to have an outside company come in and provide independent certification. So they had to go through a fairly extensive review process.



Shelley Loving-Ryder
Assistant Superintendent of the
Division of Student Assessment
& School Improvement for
the VDOE



Lan Neugent
Assistant Superintendent of the
Division of Technology Career &
Adult Education for the VDOE



Sarah Susbury
Director of the Office of Test
Administration, Scoring & Reporting
for the VDOE

Once we knew a school was ready, then just before the test, we had something we called the 96-hour checklist. That helped us a lot, because without that, we would have just been taking a chance on children's futures—having them take these tests without being positive that everything was going to work correctly. Sometimes there were problems, but none was insurmountable, thanks to the certification program.

Effects of change

Shelley Loving-Ryder: Moving to online testing requires an entirely different way of thinking about things. Now that we've completed the transition, I feel that it's probably the best thing Virginia ever did. It's provided so many benefits for us and we're just beginning to be able to completely utilize the power of being online.

I think in general our students love online testing because most of their interaction these days is with a computer or a hand-held device. In the beginning, teachers were somewhat more resistant. They were concerned about students' abilities to interact with an online assessment. And I think one of the things we quickly learned was, it was often an adult's problem. The students had no issues with taking a test online, but the teachers were concerned! Now our teachers are very appreciative of online testing, particularly because it allows them in most cases to get results back very quickly.

Parents were somewhat like teachers, in that there was some resistance in the beginning. But over time, that's become less of an issue. And again, they like the fact that they're getting results back more quickly.

Meanwhile, those assigned to set up testing sessions very much appreciate that we've gone online. The burden on them administratively has been greatly reduced. Some of our large districts actually had warehouses that were dedicated to keeping test materials! Now we have none of that.

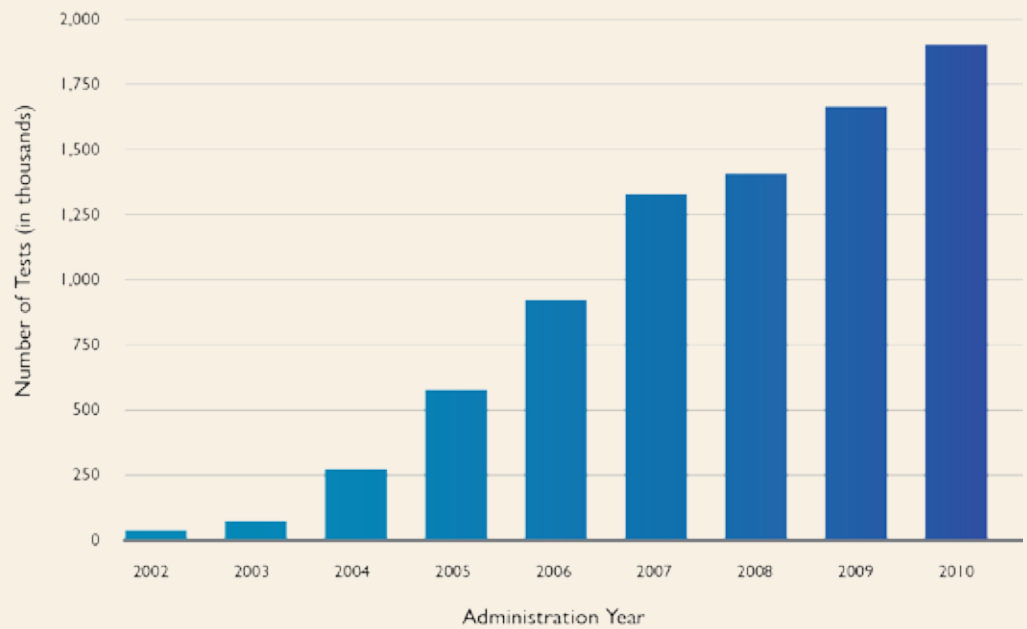
Lessons learned

Shelley Loving-Ryder: We started by taking an existing test and putting it online. That made the transition achievable and something that could be successful quickly, allowing us to gain some momentum. From here, we have the luxury of really exploring some of the technology and getting to that spot where we can do some really cool things.

A major benefit of the phased approach is that it gave us a chance to determine where we might have issues when we were still dealing with fairly small-scale testing. During this time, we worked very closely with Pearson in identifying issues that we needed to think about.

Another lesson we learned is that online testing requires a different level of support. So we

Number of Standards of Learning Tests Administered Online



learned that we needed to provide much more real-time support at the State Department level, and also realized quickly that the help desk support we needed from Pearson was somewhat different: If someone had a normal problem, then Level 1 could help them, but if there was a technology problem that was more serious, then the Level 2 people needed to be there to assist.

Sarah Susbury: To keep morale and enthusiasm high and anxiety and resentment low, we did not pressure schools to move forward with online testing. We provided some templates that gave them the important points but didn't box them into a particular plan.

Shelley Loving-Ryder: Schools had to be capable of doing online testing by particular dates, but it was never mandated that, beginning at this point, we will move completely to online testing. But once we had a couple of school divisions participate in online testing and they talked about how wonderful it was, we had no difficulty getting others to participate.

Sarah Susbury: Once districts were on board with the project, we asked them to do the same type of planning that we had done. We asked them, "What is going to be your implementation plan? Which schools are you going to transition first? Are you going to do all of your high schools at first, or are you going to just try one? Are you going to do all of the tests that we have available, or are you going to try a portion of them with a portion of students?"

The education community must come to terms with the fact that this is the way kids are learning now—I think you'll find they're more comfortable in an electronic-type environment.



CASE IN POINT: HOW MISSISSIPPI BUILT SUCCESS AND STABILITY



James Mason
Director of Student Assessment,
MDE



Jan Kirkland Hogue
Contract and Assessment Analyst,
MDE

The state of Mississippi has been using online testing to re-test high school students for their four graduation exams in Algebra I, English II, Biology I, and US History since 2003. Mississippi's online testing program began in 2001 with pilot testing, and has since grown to deliver an average of 10,000 to 20,000 tests each year across all 152 school districts in the state.

Here, James Mason, Director of Student Assessment, MDE, and Jan Kirkland Hogue, Contract and Assessment Analyst, MDE, discuss (via an interview) the challenges and successes of their state's transition to online testing and what they anticipate to be the future of online testing in Mississippi.

James Mason: I think it is important to point out that we've been through two phases. We had one or two early—I would use the word “catastrophic”—failures from the district perspective. *[Laughs]* But it was just a shared misery until we got to the current platform that we're now using, which has been much more stable, much more robust, and a much better experience.

We've gotten to the point now where online testing is as it should be, and where the technology doesn't get in the way. It has been very stable now for the past three or four years.

Stakeholder reactions

Jan Kirkland Hogue: A lot can be said about this. Students are very engaged in online testing, for the most part. They like it. The district test coordinators love it. It's less paperwork. They don't have to worry about inventorying the test materials or losing things, so they like it. They love getting the test scores back sooner.

We have had parents say students do better on a paper and pencil test. We have had parents say they do better on the computer. There are mixed reactions. But of course all parents like to get the scores back sooner. So you've got some positives, you've got some negatives. But overall when the technology works, it's a dream. When something goes bad, it's like what James said: it's a catastrophic nightmare. We have had flawless administrations for the past several years. It's either great, or it's horrible.

James Mason: The accommodation piece is the biggest thing that we have never fully exploited or gotten our hands around in online testing. That is another little logistical piece. Because right now when you do read-aloud testing, in some cases they have books, but in some cases the test administrator has to read over the shoulder of a child, which is not a very conducive test environment. So [how to address] students with special needs is a big concern.

Lessons learned

Jan Kirkland Hogue: I have a list of them:

- Testing the system before the actual test is absolutely a must.
- Training for all new users, or users who have a little experience but are still kind of rusty—that is key.
- Always have an emergency plan. The testing office in charge of this should have an emergency protocol with the testing vendor, or the sub-contractor for the online testing. If anything goes wrong that needs to be handled, you have a plan, a call-in number that you can immediately talk to the people who know what's going on, so you can formulate your specific emergency plans and communicate to your state's districts as soon as possible.
- I would say the biggest thing, each day of the test (early before it starts) is to make sure that the people in charge of the bandwidth check it, and ensure it is exactly right. So I would be very careful that all morning checks are done and rechecked. That was huge.

James Mason: A lot of lessons learned related to IT: You must think about your lowest common denominator. Right now we have some school districts with a 100-megabit pipe that is absolutely clogged, and then we have some districts with a three megabit connection that's not fully utilized. So it's not just what do you have access to, but rather the appropriate question is what are your requirements? From the company perspective, it's inexcusable in today's environment for vendors to roll out platforms on which they can't support certain capacities, or that have not been through rigorous load capacity testing. So I think the tolerance level for vendors is going to be very, very small from the state perspective.

For a paper test administration, we had always put two people in a room to administer a test—a test administrator and a test proctor. But what we found is that when you're doing online testing, you need three! Because you had to have a couple of people walking around helping and interacting and observing, but then you had to have somebody watching the computer monitor to make sure that the kids were on task, and they were working at the right place.

The greatest challenge

James Mason: I don't think the human part is as problematic now as it was. And every year that goes by, that gets easier and easier. I think it's a budgetary thing just in having the equipment or the devices to move towards more online testing. But more so than that, I think also related to budget is infrastructure.

I think it's going to be fascinating to see how vendors respond to these RFPs with the different consortiums, because I'm not convinced that the device with which you're going to be able to test 31 million children is on the street today. I don't know if it's going to be some type of tablet device ... it could be a variety of things. For instance: we could have a tablet that would come to you preloaded with the assessments and all supporting files, and then the students can take the test on this tablet/device, and when finished, you just upload the student's responses. That type scenario takes dramatically less infrastructure and bandwidth, but it's a model we have not yet fully developed or tested. You're never going to be able to test that many children on anything just online very well, using existing technology given the applications and hardware that I have seen and used over the years.

Anticipating change

James Mason: It's going to be different... I don't think we're using technology as fully as we can. One of the areas that I would be really interested in exploring is how can we leverage technology with read-aloud accommodations. Providing accommodations in testing today is very, very labor intensive, but it would not be so labor intensive in the technology-assisted environment. You could put ten different students with ten different tests in one computer lab or classroom, and just have a couple of people supervising them, whereas before it would take me 20 people to do the same thing.

My hesitancy now is on capacity. Now that our curriculum is much more rigorous with standards that are more in line with national assessments like NAEP, the failure rates have gone up significantly and dramatically—to the point that we wouldn't have the technology base in all of our schools to support a re-test with our new curriculum framework.

The n-count for the online tests under the old curriculum is now very low as students are graduating or just moving on. My main concern is having the capacity to test the large number of students who do not pass the more rigorous tests. If you really have a rigorous assessment system, are you going to have the capacity in place to deal with it in an ongoing basis?

Moving forward

James Mason: I think people need to be very deliberate both at the user end, and at the vendor level, in pushing change too quickly. Again, what you've got now might be great, but the fatigue factor with change can take an incredible toll, even when the change is for the better.

I also think it is important that we make distinctions between instructional technology and assessment technology. Someone recently pointed out this distinction at an EIMAC meeting and I think it is a critical discussion that needs to take place. With the proliferation of tablets, smart phones and other devices, our constituencies will likely not understand why we are not able to change more quickly in the area of large scale assessments.

Jan Kirkland Hogue: I would recommend doing a small pilot in any state who is transitioning to an online system or even a new online system to work out the kinks. You need to expect the unexpected. You cannot imagine the different things that can come up with online testing. For instance, the very first time that we used it, we had a meteorite storm that completely shut down the whole system. Now, who would have ever thought that would have happened. Such a storm hadn't happened in 400 years, and on the very first day of testing, that's what happened and we couldn't finish testing that day. So I would say expect the unexpected.



CASE IN POINT: BRIDGING TECHNOLOGY & ASSESSMENT IN NORTH CAROLINA

The North Carolina Department of Public Instruction (NCDPI) began conducting online tests in 2004–2005 to assess achievement in the computer skills curriculum. End-of-course exams followed, beginning in 2006–2007 with Physics. Currently, the state is field-testing online tests for the North Carolina Essential Standards and the Common Core State Standards, with the goal of conducting all state assessments online by the 2014–2015 school year.

Here, representatives from a cross-section of the North Carolina student assessment program discuss the state’s transition to online testing: Jim Kroening, Lead Testing/Accountability Consultant, NCDPI; Randy Craven, IT Director at NC State-TOPS; Kayla Siler, Policy and Strategic Planning, NCDPI; and Nadine McBride, Psychometrician, NCDPI.

Utilizing studies and surveys

Kayla Siler: We surveyed our LEAs and charter schools about their readiness to transition to online assessments. The survey focused on readiness in terms of hardware, connectivity, and personnel. The survey questions included questions about number of devices, devices for student use, mobile devices or labs, wireless connectivity, bandwidth, and personnel available at the school or district level to assist with technology issues.

We had pretty good response rate, and completed our first online readiness report in May 2010. And then, the addendum that came out earlier this year (February 2011) included some additional responses that were received over the summer and into the fall of last year.

As part of a separate effort, we also interviewed about 20 LEAs or charter schools in face-to-face interviews about the actual use of online assessment in their schools.

Nadine McBride: A couple of years ago, when we were just beginning to consider moving the modified alternate assessments online, we conducted an online feasibility study. And we learned a lot from it. This year, we conducted an online item tryout of the entire population assessed on the modified alternate assessment. With it we conducted a student survey in order to get student input. We’ll be analyzing that data over the next week or two. We expect to learn a lot from that as well and to use the result to assess if we need make any kind of adjustments.

We’re also looking at conducting some special studies of the technology-enhanced items, of [doing] an in-depth usability study of those technology-enhanced items with the younger kids, just to make sure they’re interacting with the items in the way that we expect them too.



From left to right: Jim Kroening, Lead Testing/Accountability Consultant, NCDPI; Kayla Siler, Policy and Strategic Planning, NCDPI; Nadine McBride, Psychometrician, NCDPI; Randy Craven, IT Director at NC State-TOPS

Jim Kroening: We also conducted a test administrator survey following this year's item tryout. Some of the questions were very similar to what we asked the students. We are going to be looking at the perceptions of test administrators, versus what the thoughts of the students actually were. The data should be interesting because those two things at times can be very different.

Kayla Siler: We're going to do an annual update at the beginning of the school year, to let the NC State Board of Education know where we are, on track or not, with meeting online readiness goals. This fall our update will say, here's an update of what happened in the '10-'11 school year, and here's what you should expect for next school year, and the next school year, as we're looking towards our goal of being all online by '14-'15.

Meeting testing load demands

Randy Craven: When we started development of our first online assessment, the online test of computer skills, our biggest concern was being able to spread the load, both locally and on the hosted side, across a wide window, so that testing would not be attempted on the same day with 100,000 students, all starting at 7:30 a.m. on the same morning.

Policy was established to open up the testing window to be basically all semester long, both fall and spring semesters. Local districts had to schedule their sessions based on [a] centralized management tool that we provided which kept the number of concurrent testers to a set maximum for valid testing times. Load management was critical to the success of the project. With the

cumulative or spiraled design of the curriculum for computer skills, it made sense that students shouldn't have to wait until the end of a semester or a year, to take the assessment.

That's in stark contrast to what folks want to do for an end of year or an end of course sequence. They don't want to test before the end of the series, and don't want to test too many days after the end of the series. So, there's a very limited range on what folks really want to do for most online assessments. They would like to test a million students on one day in North Carolina. And that puts a big burden on trying to design a system that can accommodate that.

Jim Kroening: It is going to be difficult to promote expanded testing windows and more flexibility to our school clients. When dealing with end-of-course tests or other types of summative tests, folks don't want to test students early in the window because they want to get as much instructional time in as possible.

Adjusting strategy with growth

Jim Kroening: We are more aware of the importance of marketing and promoting online testing now than we were in 2004-2005. In some ways, I believe we really haven't evolved as fast as we should have into an online environment for both instruction and assessment.

Promoting the benefits of online testing cannot be understated. Our online testing interface is user-friendly making the tests as accessible as possible for as many students as is possible.

For example, read-aloud accommodations require special settings, test administrators, and proctors. If students who need this accommodation test online with headphones, a lot of staffing issues that occur at local schools during testing times would be alleviated.

Kayla Siler: We found that the planning side is what people often overlooked. They thought, we'll just do it online. We can do that anytime. They didn't think about, [how] that means [they] actually have to have 30 laptops in the room, so that all these students can take a test on this day at this time that we signed up for. And they all have to be connected to the Internet, and they all have to have power, and they all have these needs. They didn't take the planning time to say, okay, this many kids in this course are taking this test, on this day, at this time, in these places. That's where they ran into problems that made them not want to do online assessment again.

Other LEAs took the time to plan and now they're really building on a system. Some of our districts with the most success took a couple of years to see what it would look like, tried one class, tried one course, and built up to the point where, now, it's very easy for them to administer the assessments online. Scheduling is no problem. They go in early and get all the spots they need to register for the test. They have all their students registered. And they're testing 60 or 80 percent of their population

online with no trouble. Now it's just getting everyone to take that approach or to at least give online assessments a try, rather than stick with the old, and what they know, and what they're comfortable with.

Nadine McBride: We're trying to promote the fact that the online assessment has more capability to have built-in accommodations. For example, it will make read-alouds a lot easier for the schools. We're really trying to promote that accessibility and not just for the modified population, but also for the grade five and eight science students. We are also promoting the interactivity of the technology-enhanced items that we can have with online assessment that they'll be more interesting and engaging to the younger students.

Randy Craven: On the technology side, we are making some architectural adjustments, and have plans to try to offload some of the delivery mechanisms to other systems, so that the data collection side can catch responses securely. And we're trying to figure out ways to alter our architecture, so that we can increase our performance about tenfold [in order] to have the capacity to test close to a million students in a day.

Lessons learned

Kayla Siler: What we learned from talking with our districts about some of the keys [to successful online testing was that] the number one for most of them was the collaboration between testing personnel and technology personnel. And so, I think, at the school level, at the district level, and even at the state level, that that is an important piece, that you have both teams working together to make it successful, that you know who to contact when there's a problem, that you know who's responsible for what, and that you work around scheduling together. So if you're sharing computers across schools, or across classrooms, you know when you can have them where, and how to set them up, and who's responsible for what.

After that, I think it was communications, not only between those groups, but [also] with the teachers in the school. [It was important to make] sure that the teachers were aware of what their options were, that they had the training that they needed to be able to administer the test online or paper/pencil, and [that] everyone in the school knows what time the testing is happening, where it's taking place, and where the students are supposed to be, so you get people in the right places.

I do think that when superintendents and other district leadership were supportive, that made a difference, because then the staff felt that they were accountable to do what the superintendent was asking. If [the leadership] weren't pushing [online testing], there was less accountability, or less incentive to make that change.

Resources and final words

Kayla Siler: On our NC Public Schools website (<http://www.ncpublicschools.org>), we do have a space where we've tried to post resources [for online testing]. Some of that has been just looking for education-related articles about the use of technology in the classroom for assessing, for different things, about formative and benchmark assessments, about the different types. We also reference other states that have online systems and maybe have FAQs or different documents that could be applied anywhere.

Randy Craven: I have always felt that it is very important to have a technical staff that is aware of accountability issues and processes, not just a relationship with IT support that is just about technology.

Jim Kroening: It is important to consider where the vision comes from and who promotes the vision. During the early years when we developed the online test of computer skills, the test development staff promoted the initiative. It quickly became a State Board of Education initiative and a graduation requirement. Today, the vision originates from the leadership as part of the Accountability and Curriculum Reform Effort (ACRE) and Race to the Top (RttT) programs. Since online learning and testing is a statewide initiative, it can be more effectively promoted and supported by our school clients.



Copyright © 2011 Pearson Education, Inc. or its affiliate(s). All rights reserved.