

Recommendations for Deploying Innovative Technologies to Create Better Assessments (Question #1)

Randy Bennett
ETS
rbennett@ets.org

Presentation at the Race to the Top Assessment Program
Public & Expert Input Meetings, November 13, 2009, Boston, MA

Question #1

- *Propose how you would recommend that different innovative technologies be deployed to create better assessments, and why. Please include illustrative examples in areas such as novel item types, constructed response scoring solutions, uses of mobile computing devices, and so on.*

Recommendation #1

- Start with a long-term vision (5-10 years out) for a next-generation assessment system
 - Work backward to a set of steps to get there, including significant near term ones
- Why?
 - It takes 2-3 years to create, review, field test, calibrate and administer a new parallel form of a *paper-and-pencil multiple-choice* test
 - If 3-4 years is the end-state time frame, the likelihood of achieving fundamental change is *not* great

Recommendation #2

- In that long-term vision (and to the extent possible in the incremental steps), focus on such critical ideals as using technology to:
 - a. Measure important competencies that can't be measured well in the conventional form
 - b. Help teachers (and students) adjust instruction and learning
 - c. Model effective teaching and learning practice
 - d. Use technology to make assessment fairer for all students, including those with disabilities and English Language Learners
- Why?
 - Worthwhile *near-term* efficiency targets (e.g., improving score turn-around) may dominate to the detriment of more fundamental goals (e.g., measuring what's important)

Recommendation #2a

- Use technology to measure important competencies that can't be measured well in conventional form, e.g., having students:
 - Use simulations of dynamic systems to interpret evidence, discover relationships, infer causes, pose solutions
 - Mathematically model problem situations with a spreadsheet
 - Write on computer and read (nonlinearly) on the Internet
 - Search for, and critically evaluate, information on the Internet
 - Respond to reading or writing problems that require the integration of many text sources and of various document types
 - Fluently execute basic procedures (formatively useful)
 - Carry out complex extended projects
 - Assemble digital portfolios of their work

Recommendation #2b

- Use technology to help teachers (and students) adjust instruction and learning, e.g.,
 - When a student’s summative test performance suggests the presence of either an overall proficiency deficit or of specific skill deficits:
 - At the least, provide “formative hypotheses” that point teachers toward students or skill areas of need, upon which teachers (and students) should follow up
 - Route the student to a targeted diagnostic assessment

Recommendation #2c

- Use technology to model effective teaching and learning practice, e.g.,
 - Build tools into test questions that practitioners use, and that students should be using routinely, in the course of their domain practice, e.g.,
 - Planning tools for writing
 - Graphical organizers and tables for representing complex text (with appropriate alternatives for students with visual disabilities)
 - Concept maps for representing physical or semantic relationships

Recommendation #2d

- Use technology to make assessment fairer for all students, including those with disabilities and English Language Learners, e.g.,
 - Vocabulary links for difficult words (where vocabulary knowledge is not being tested)
 - Alternate representations of the same information (text to speech, described graphics)
 - Alternate questions measuring similar skills at similar difficulty levels, when a class of questions is important but not suitable for some students

Recommendation #3

- Understand the benefits and limitations of each technology before deploying, e.g.,
 - Automated scoring
 - Operationally faster and cheaper than human scoring, and sometimes able to provide feedback on instructionally actionable performance components
 - In many cases, uses limited proxy measures to *predict* a human score; practicing the proxies may lead to higher machine scores but not necessarily to greater skill
 - Adaptive testing
 - Measures with precision throughout the skill range
 - In current implementations, measures only a subset of what's important to test, potentially having the same (unwelcome) effects on instruction as current state assessments are said to have
- Why?
 - There are tradeoffs associated with new technology that are best made by informed choice, rather than by accident

Recommendation #4

- Manage risk
 - Most successful transitions from paper-and-pencil to computer delivery have put substantial time into planning and many have used a phased approach to implementation
 - Examples: Oregon, Virginia
- Why?
 - Moving a large-scale testing program to computer is a *very* complex undertaking
 - Hardware and software availability and compatibility in all schools
 - LEA Training
 - Student familiarization

Recommendation #5

- In the world of innovation, failure is a fact of life but one that can be put to beneficial use, so *plan to fail*:
 - Early
 - Often
 - Small
 - Gracefully
- Why?
 - We'll either learn relatively quickly that an approach is unworkable or, in the best case, successively approximate over time a practical assessment system with the least cost and harm to all concerned

Recommendation #6

- Fund multiple consortia so that significantly different assessment models (and uses of technology) can be explored and compared to one another
 - Consider giving preference to models that already have an existing theoretical base and have been piloted
- Why?
 - We know a lot about how to create innovative (technology-based) assessments, so we should build on that existing knowledge
 - We know a lot less about how to create innovative (technology-based) assessments that are affordable, practical, technically defensible, accessible, and fair to all students, so there is great value in funding multiple approaches

Summary of Recommendations

1. Start with a long-term vision (5-10 years out) for a next-generation assessment system
2. Focus on such critical ideals as using technology to:
 - a. measure important competencies that can't be measured well in conventional form
 - b. help teachers (and students) adjust instruction and learning
 - c. model effective teaching and learning practice
 - d. make assessment fairer for all students, including those with disabilities and English Language Learners
3. Understand the benefits and limitations of each technology before deploying
4. Manage risk
5. Plan to fail early, often, small, and gracefully
6. Fund multiple consortia so that significantly different assessment models (and uses of technology) can be explored and compared to one another