



NCES Initiative on the Future of NAEP

Edward Haertel, Panel Chair
Stanford University

New Orleans, LA

March 2, 2012

Panel Membership

- Edward Haertel (*Chair, Stanford University*)
- Russell Beauregard (*Intel Corporation*)
- Jere Confrey (*North Carolina State University*)
- Louis Gomez (*University of California, Los Angeles*)
- Brian Gong (*National Center for the Improvement of Educational Assessment*)
- Andrew Ho (*Harvard University*)
- Paul Horwitz (*Concord Consortium*)
- Brian Junker (*Carnegie Mellon University*)
- Roy Pea (*Stanford University*)
- Bob Rothman (*Alliance for Excellent Education*)
- Lorrie Shepard (*University of Colorado at Boulder*)

Charge to the Panel

- **Charge:**

Develop a high-level vision for the future of NAEP as well as a plan for moving NAEP toward that vision

- **Audiences:**

NCES, NAGB, NAEP Contractors, Policy Makers, Researchers, Concerned Citizens

Timeline

- NAEP Summit
- Panel Convened
- Draft Outline
- NAEP Summit (SEAs)
- Final Outline
- Form writing groups
- Initial Draft
- Final Draft
- Deliverable
- August 18-19, 2011
- October 2011
- November 2011
- January 24-25, 2012
- January 2012
- January 2012
- February 2012
- March 2012
- March 31, 2012

A Work in Progress...

- *We are still discussing more ideas than are likely to appear in our March 31 draft report. What I am presenting today reflects some of our current thinking, but we have not reached consensus on all of these ideas. There is no assurance that any particular idea will appear in the final version.*
- **PLEASE SHARE YOUR REACTIONS AS WELL AS YOUR OWN IDEAS!**



Overview of today's presentation

- Context for this initiative
 - NAEP infrastructure
 - NAEP content frameworks
 - NAEP and technology
 - Embedded assessments
 - Reporting
- NAEP's continuing importance



Context for This Initiative

- A Changing Environment for NAEP
- More Ambitious Expectations
- Rapid Change in Technology

A Changing Environment for NAEP

- Preparing students for a changing world
- Relation to the CCSS
- Relation to PARCC and SBAC assessments
- Globalization
 - State participation in TIMSS, PISA, ...
- Evolution of “education” beyond “schooling”

More Ambitious Expectations

- Reasoning and problem solving in complex, dynamic environments
- Communication and collaboration
 - Group problem solving
- Expanded views of literacy
 - Identifying need for, locating, and evaluating information
 - Fluency with new technologies (e.g., TEL)
- College and career readiness

Rapid Change in Technology

- Increasing educational use of (e.g.):
 - e-textbooks
 - interactive media
 - web-based resources
- Increasing availability of
 - massive data warehouses
 - data mining
- Increasing communication/cooperation as states move toward "shared learning infrastructure"



NAEP Infrastructure

- Background
- NAEP's place in "Assessment Ecosystem"
- NAEP Innovations Laboratory?
 - Illustrative topics

Background

- NAEP is a Complex System
 - Involves multiple organizations, areas of expertise
- R&D Component is critical
 - NAEP not only tracks achievement, but also drives assessment innovation.
 - NAEP's methodology is emulated worldwide.
 - NAEP R&D is guided and funded through multiple, complex institutional mechanisms.
 - Systematic review might identify possible improvements.
- NAEP as backbone of “assessment ecosystem”?
- NAEP Innovations Laboratory?

Evolving Assessment “Ecosystem”

- Potential role for NAEP as backbone of evolving assessment infrastructure
 - Design changes to facilitate linkages between NAEP and other assessments
 - Bridge between state-level assessments and TIMSS, PISA, PIRLS, ... ?
 - Explicit attention to NAEP vis-à-vis the CCSS
- Defining the state-of-the-art in large-scale assessment

NAEP Innovations Laboratory?

- Purposes
 - Strengthen and systematize NAEP R&D
 - Strengthen linkages to other assessment programs and facilitate dissemination
- Features
 - Access point for vetting new ideas
 - Organizational structure not yet specified
 - Would support both in-house and 3rd party studies

NAEP Innovations Laboratory?

- Step 1: Review existing structures for NAEP R&D
 - Design and Analysis Committee
 - Validity Studies Panel
 - IES Program on Statistical and Research Methodology in Education
 - NAEP Data Analysis and Reporting contract
 - Education Statistics Support Institute Network (ESSIN)
 - NAEP Secondary Analysis Grants program
 - ...

NAEP Innovations Laboratory?

- Step 2: More clearly frame purposes NAEP R&D should serve
 - investigate / assure validity of NAEP findings
 - improve NAEP processes to reduce testing time, reporting time, measurement error, cost
 - expand the range of constructs assessed
 - enable NAEP to serve new purposes
 - e.g., linking to other assessments

Illustrative Topics

- Assessing home-schooled students? lower grades or pre-K? college students?
- R&D on new item types
- Interpretability of NAEP reports
- dynamic (evolving, non-static) content frameworks
- adaptive testing
- technology-enhanced accommodations
- Linkage to state data systems
- Linkage to massive data warehouses
- ...



NAEP Content Frameworks

- Relation of NAEP content frameworks to the Common Core State Standards
- Dynamic content frameworks?

Relation to the CCSS

- Considered and rejected:
 - “CCSS” as replacement for NAEP frameworks
 - “CCSS” scales within NAEP
- Distinct functions for NAEP vs. CCSS
 - NAEP covers more content areas
 - NAEP covers more content within each area
 - in part due to focus on populations, not individuals
 - Broader frameworks facilitate linking
 - Value in multiple content frameworks
 - Measuring content outside the focus of instruction can inform wise deliberation regarding evolution of C&I

Dynamic Content Frameworks?

- **Current approach:**

Content frameworks are held constant for a period of time, then replaced

- **Alternative to consider:**

- Framework development panels are replaced by standing committees of content experts
- Achievement is defined relative to a mix of knowledge and skills that is updated incrementally, analogous to the CPI
 - *Affords local continuity, but broad constructs may evolve over time*
 - *Caution: Implies need for a principled way to place relative values on alternative learning outcomes*



NAEP and Technology

- Technology and teaching-and-learning
- Technology and assessment
- Deeper links to state data systems and massive data warehouses

Technology & Teaching-&-Learning

- How might NAEP monitor the full range of complex activities students pursue in modern learning environments?
 - Changing representational modalities and user interface modalities
 - Gesture and touch, sketching, voice, visual recognition and search, augmented reality
 - Interaction with dynamic knowledge representations

Technology and Assessment

- Assessing Old Constructs in New Ways
 - New platforms for item presentation
 - New modalities for student response
 - Adaptive testing to improve precision (especially in the tails of the distribution)
- Assessing New Constructs
 - Technology and Engineering Literacy
 - Problem solving in adaptive environments
 - Technology-mediated collaboration

Deeper Links to State Data Systems

- Expand on current use of state data to improve efficiency of within-state samples
- Expand on initial efforts linking NAEP scales to State assessment scales
 - E.g., mapping of state “Proficient” definitions to NAEP score scales
- Consider building and maintaining integrated longitudinal data structures
 - Interpreting student performance on new kinds of tasks may require knowledge of instructional history



Embedded Assessments

- Embedded Assessments?
- Why EAs in NAEP?
- Can NAEP bridge the gap?

Embedded Assessments?

- Unclear exactly what EAs are, but most accounts suggest:
 - Assessments are linked more closely to ongoing instruction
 - Students engage in complex tasks or create complex (scorable) work products
 - Problems may have multiple solutions
 - Data collection may be “transparent” or unobtrusive
 - Standardization is much weaker than for conventional assessments (task specifications, administration conditions, scoring rubrics)

Why EAs in NAEP?

- Fundamental challenge of standardized assessment where there is no standard curriculum
 - Each item must be a self-contained package, providing all relevant context and content
 - Interpretation of test performance may be unclear if instructional history is not known
- Test tasks must be brief
 - No writing revision cycles, for example

Can NAEP bridge the gap?

- “EA” reflects the perennial desire to link classroom and external assessments
- May not comport with structure and culture of US educational system

Reporting

- Revisit role of achievement levels
- Greater use of “active” reporting
- Reporting R&D

Revisit Role of Achievement Levels

- Achievement Levels are popular
 - ALs are the major vehicle for NAEP reporting
 - ALs respond to demand for normative information
 - ALs seem easy to understand
 - AL reporting is widely emulated and written into law
- Achievement Levels are problematical
 - ALs are poor choice of statistics to summarize distributions
 - ALs are prone to serious misinterpretation
 - ALs are invested with surplus meaning
- Revisit heavy reliance on ALs for NAEP reporting
 - No recommendation to abolish ALs, but instead to accompany them with better methods for meaningful reporting

Greater Use of “Active” Reporting

- Dynamic displays, interactive graphs
 - NAEP Data Explorer and related tools are great resources
 - Bring these methods into online reports with embedded animations and customizable displays
 - These applications might be web-based or stand-alone

Reporting R&D

- Periodic audit of media accounts to identify any systematic miscommunications
- Research on usability of active displays
- Research on alternatives to Achievement Level reporting
- Research on communication about supported versus unsupported interpretations
 - Complexity will increase with availability of cross-assessment linkages

NAEP's Continuing Importance

- NAEP's traditional functions are still critically important
 - Low-stakes “audit” function
 - Designed to estimate achievement distributions, not individual scores
 - Source of assessment innovations critically needed to fulfill promise of new consortia
- NAEP may evolve to serve additional purposes
 - Backbone of new assessment “ecosystem” anchoring linkages among multiple assessment programs

Thank you