



Jim Pellegrino
Learning Sciences Research Institute
University of Illinois at Chicago

UIC <http://www.lsri.uic.edu/>



*Assessment as a Trojan Horse for
Educational Improvement & Equity*




Main Points

- Assessment can serve both positive and negative functions in educational systems
 - **Negative** when it becomes the de-facto driver of curriculum and instruction (often the product of large-scale testing & program evaluation efforts)
 - **Positive** when it breaks down barriers and forces diverse stakeholders to have a deep conversation about what we want students to know and be able to do and the implications for teaching & learning
- In the context of accountability, it is especially important to sort out these issues so that assessment can be a positive *"Trojan Horse"*


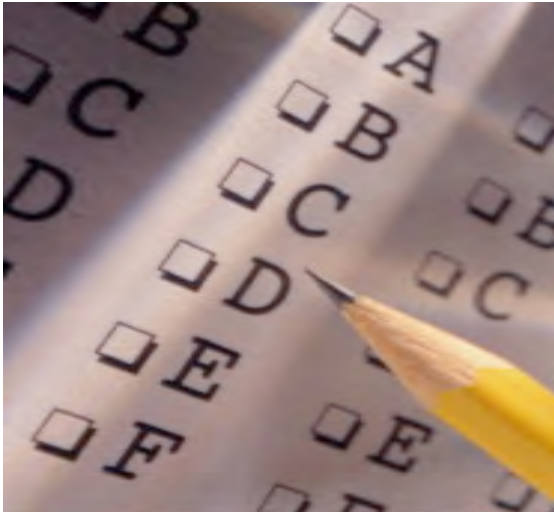


Elements in Making My Case

- When the assessment of **"X"** substitutes for understanding what **"X"** is and how it develops
 - Examples from Large-Scale testing programs

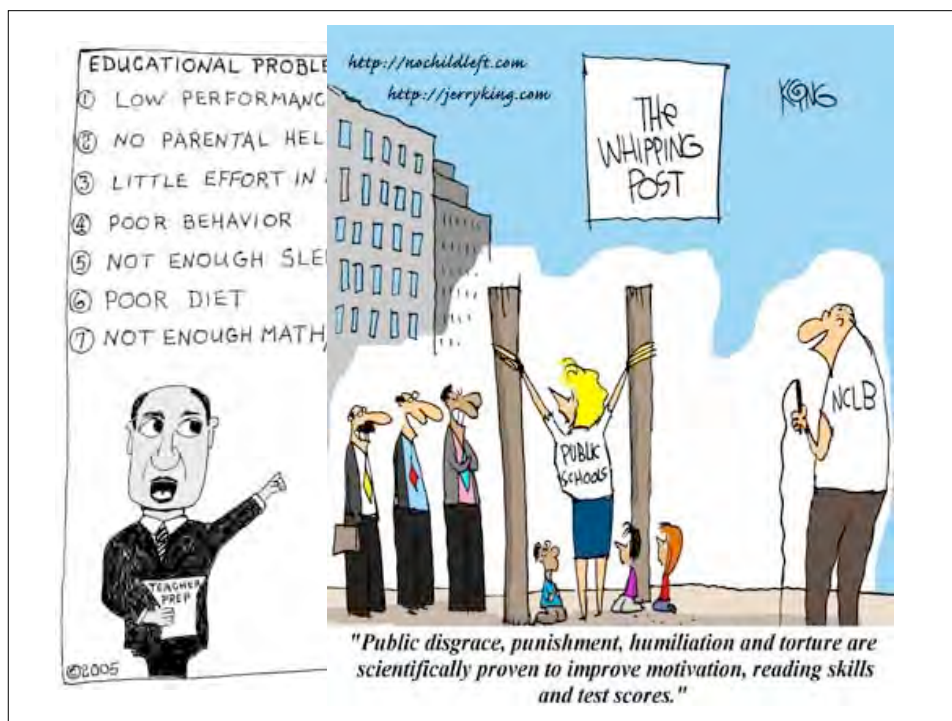


An Obsession with Standardized Testing & Test Scores



NO CHILD
LEFT ~~BEHIND~~
UNTESTED

PRESIDENT GEORGE W. BUSH



Perspectives on the Problem

- “Assessing education well may depend on assessing it less.” (Stake, 1999)
- You don’t fatten a pig by constantly weighing it!
- What we are testing is but a small part of what students are really supposed to know & understand.
- The focus on standardized tests is unduly narrowing the teaching and learning process.
- Current tests are based on outdated theories & technologies.




Elements in Making My Case

- When the assessment of “X” substitutes for understanding what “X” is and how it develops
 - Examples from Large-Scale testing programs
- The conceptual underpinnings of assessment
 - What needs to drive the thinking & design

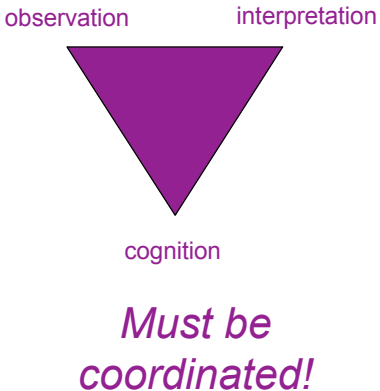
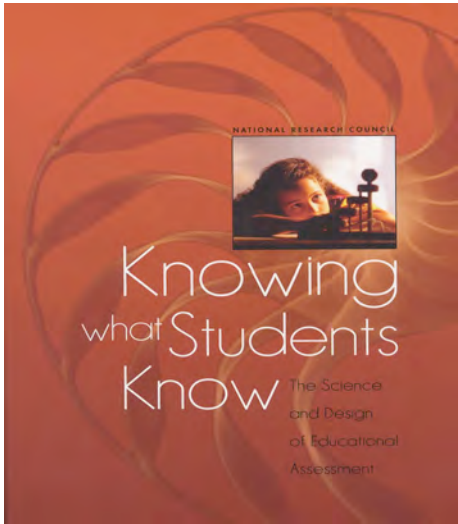


Contexts and Purposes: Distinctions That Matter

- Contexts:
 - small scale: individual classrooms
 - intermediate scale: schools, departments, colleges
 - large scale: university systems, districts, states, nations
- Purposes:
 - assist learning (formative)
 - appraise individual achievement (summative)
 - evaluate programs (accountability)
- Problem: One size does not fit all
 - Educators at different levels need different information
 - Differing priorities, constraints, & tradeoffs




Assessment as a Process of Reasoning from Evidence



observation interpretation

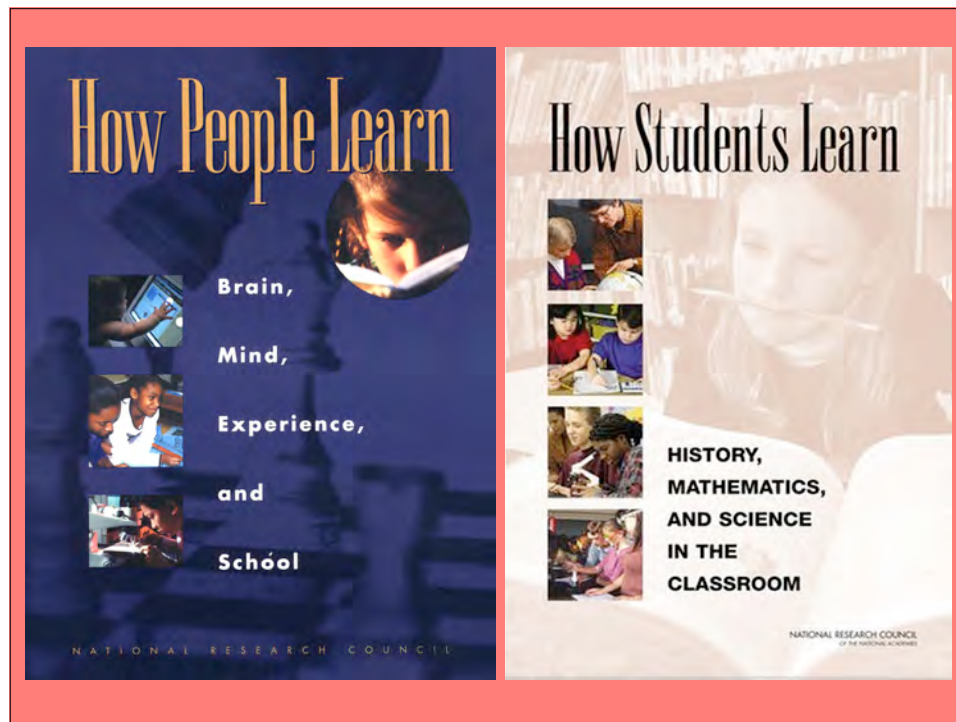
cognition

Must be coordinated!



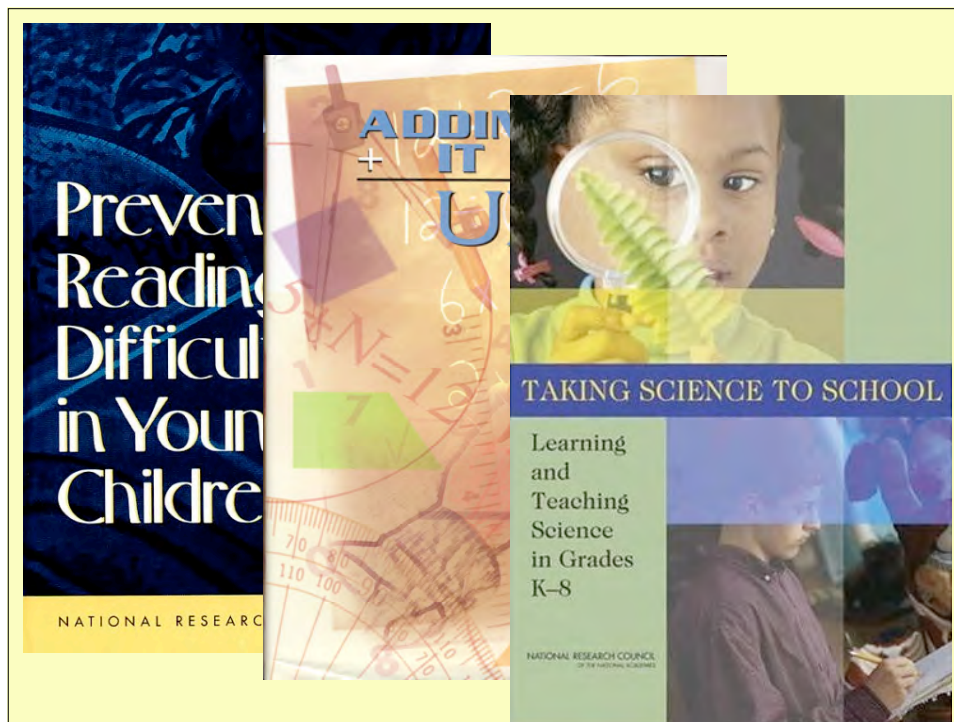
Scientific Foundations of Educational Assessment

- Advances in the Sciences of Thinking and Learning -- **the cognition vertex**
 - informs us about what observations are important and sensible to make
- Contributions of Measurement and Statistical Modeling -- **the interpretation vertex**
 - Informs us about how to make sense of the observations we have made



Why Models of Development of Knowledge & Skill are Critical

- Tell us what are the important aspects of knowledge & skill that we should be assessing.
 - much more precise and interpretable than “standards”
- Give us strong clues as to how such knowledge can be assessed
 - target “learning performances” and useful forms of evidence
- Can lead to assessments that yield more instructionally useful information
 - diagnostic & prescriptive
- Can guide the development of systems of assessments
 - work across contexts, levels, & time



Elements in Making My Case

- When the assessment of “X” substitutes for understanding what “X” is and how it develops
 - Examples from Large-Scale testing programs
- The conceptual underpinnings of assessment
 - What needs to drive the thinking & design
- The challenge of going from rhetoric to reality
 - Redesign of a major high-stakes program



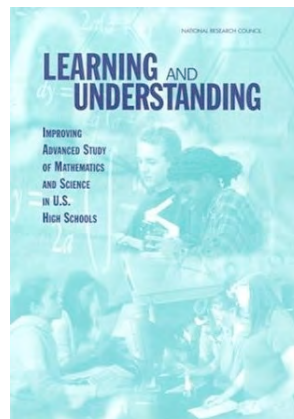
CollegeBoard Advanced Placement Program

- Since 1955, the AP Program has enabled millions of students to earn college credit or placement while still in high school.
- The AP Program offers 37 courses and exams.
- More than 17,000 schools worldwide participate in the AP Program.
- Twenty-five percent of U.S. public high school students in the class of 2008 took an AP Exam at some point in high school.
- In 2008, nearly 1.6 million students worldwide took more than 2.7 million AP Exams.
- More than 90% of the nation's 4-year colleges and universities, and institutions in 45 countries, have an AP policy that grants incoming students credit, placement or both for qualifying AP Exam grades.
- In 2008, more than 3,600 colleges and universities accepted qualifying AP Exam grades for credit and/or placement.



Why an AP Science Redesign?

- **A 2002 NRC Report identified ways to improve advanced study of math and science in the U.S. The Report's recommendations are applicable to all AP course subjects:**
 - Emphasize deep understanding rather than comprehensive coverage -- avoid "mile wide & inch deep" syndrome
 - Reflect current understanding of how students learn in a discipline
 - Reflect current research directions within the disciplines
 - Emphasize the development of inquiry and reasoning skills





AP Redesign Goals

Produce a more inclusive and more engaging program of study for each AP discipline by identifying:

- The concepts to be studied in depth and measured on the exams
- The essential reasoning and inquiry skills that are to be supported with instruction and measured on the exams
- Cutting-edge areas of research that can capture essential concepts within the discipline and engage diverse student populations
- Instructional practices that are successful for developing conceptual understanding
- The minimum resources required to support these practices

To Create an AP Program that:

- Engages diverse student populations in college-level knowledge, skills and abilities within a discipline
- Supports teachers with professional development opportunities and tools
- Supports the awarding of college credit/placement for qualifying student exam performance



Conceptual Approach Builds Upon Work of Others:

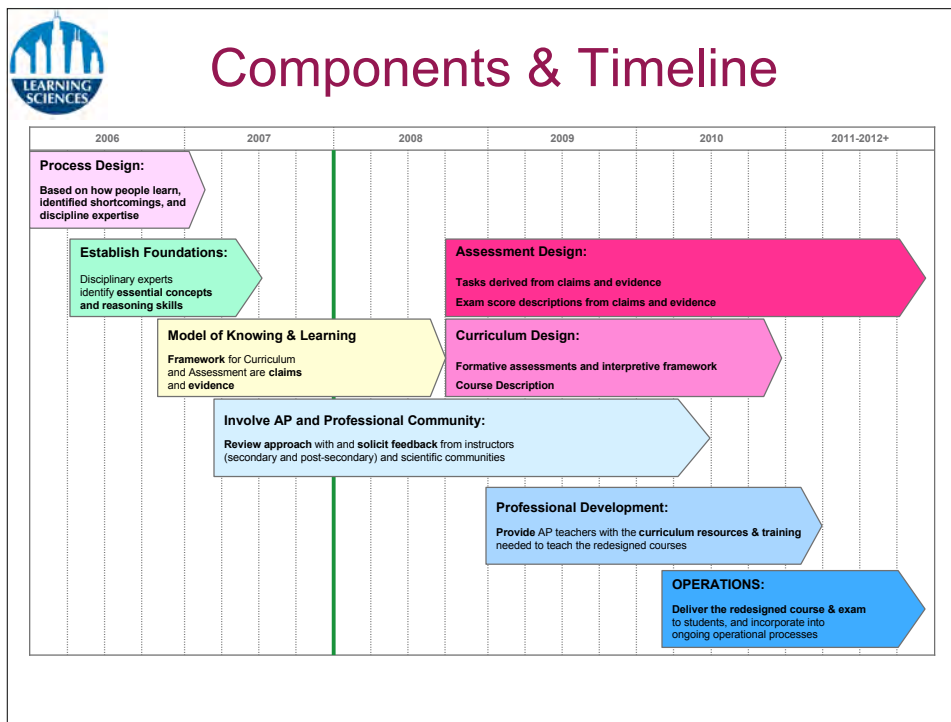
Understanding by Design, G. Wiggins & J. McTighe (1998, 2006).

How People Learn: Brain, Mind, Experience & School. J. Bransford, A. Brown, & R. Cocking, S. Donovan, & J. Pellegrino (2000).

Knowing What Students Know: The Science and Design of Educational Assessment, J. Pellegrino, N. Chudowsky, & R. Glaser (2001).

Evidence-Centered Assessment Design: Layers, Structures, and Terminology, R.J. Mislevy & M.M. Riconscente (2005).

Collaboration with Jeanne Pemberton, Mark Reckase, Meryl Bertenthal, John Eggebrecht, Kristen Huff, Cindy Hamen, Marcia Wilbur et al.
Supported by NSF and the College Board




AP Redesign Implications

For AP science teachers and students:

- AP instructors and students will have a well-defined set of learning objectives that support teaching for deeper understanding.
- The AP Exams will be congruent with these learning objectives.
- AP instructors will have tools and professional development opportunities that support teaching, learning and success on the AP Exam
- The post-secondary community and professional societies will have a better understanding of, and confidence in, the value of AP courses.

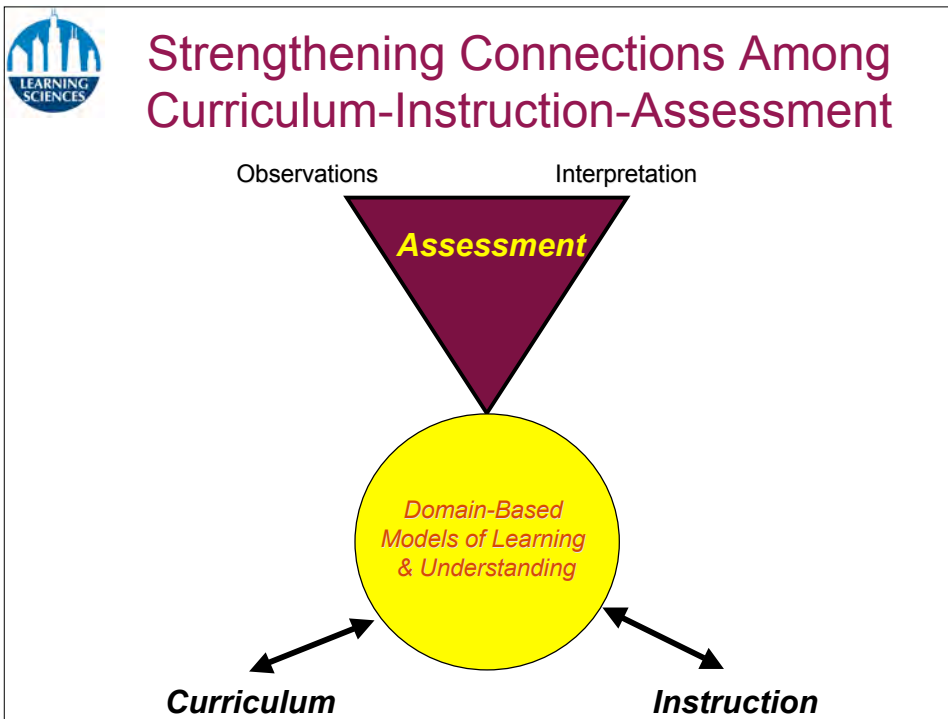
For science education generally:

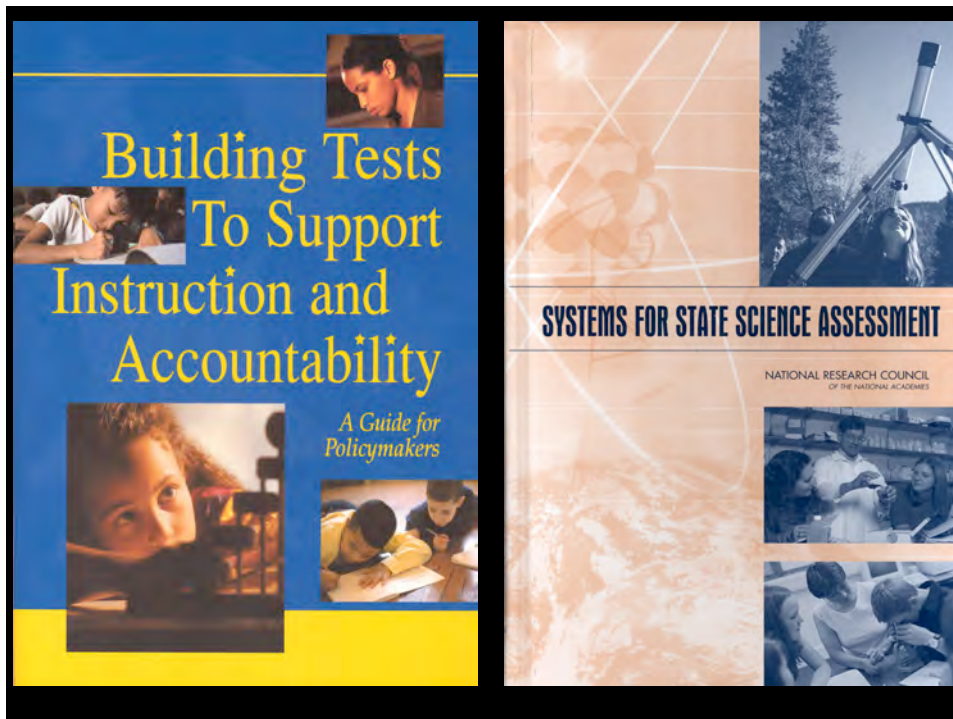
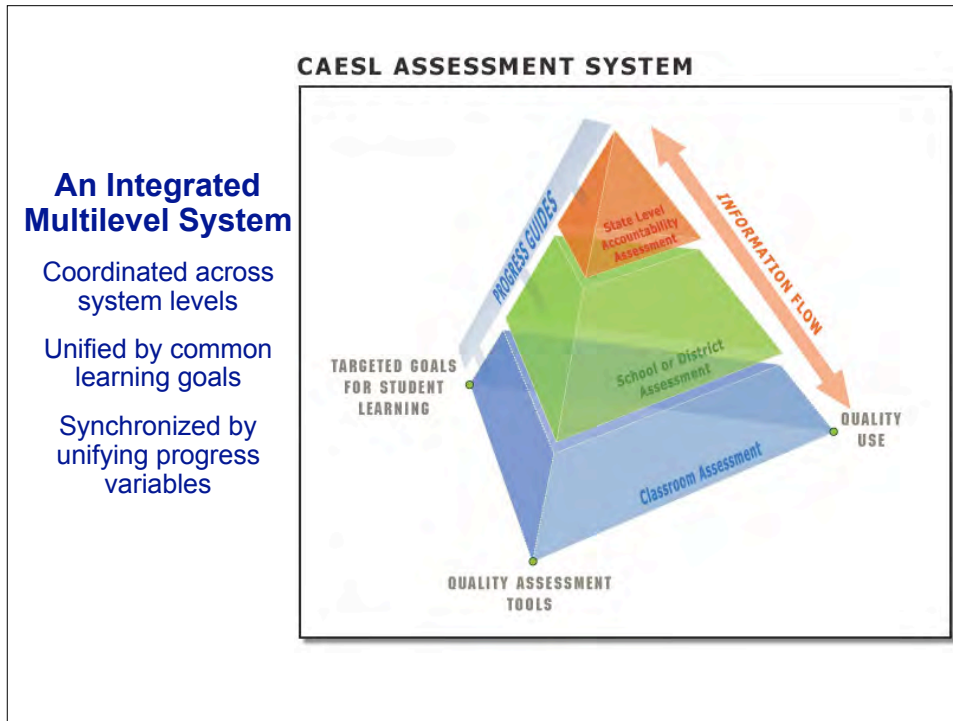
- The emphasis on reasoning and inquiry, enforced by the exams, can influence school science in lower grades and in the universities.
- The development of a high-stakes exam based on evidence-centered design principles can influence state and national assessment designs.
- The redesigned courses will increase interest and success within a new population of students who can then contribute to both science education and the practice of science.



Elements in Making My Case

- When the assessment of “X” substitutes for understanding what “X” is and how it develops
 - Examples from Large-Scale testing programs
- The conceptual underpinnings of assessment
 - What needs to drive the thinking & design
- The challenge of going from rhetoric to reality
 - Redesign of a major high-stakes program
- Applying these ideas to developing state plans for assessment & accountability
 - Building coherent systems of assessment







Elements in Making My Case

- When the assessment of “X” substitutes for understanding what “X” is and how it develops
 - Examples from Large-Scale testing programs
- The conceptual underpinnings of assessment
 - What needs to drive the thinking & design
- The challenge of going from rhetoric to reality
 - Redesign of a major high-stakes program
- Applying these ideas to developing state plans for assessment & accountability
 - Building coherent systems of assessment
- Some Final Thoughts

We need to invest in building better assessments!



**We need to learn how to help teachers use them
to improve student learning!**



What Will Be Needed to *Leave No Student or Teacher Behind*

- Instructional & assessment materials that incorporate knowledge about domain specific learning trajectories in areas of math, science, literacy, social science etc.
 - Not using what we currently know; need to know more
- Better balance and coordination between large-scale & classroom/instructional assessment practices
- Teacher education -- pre-service & prof development
 - Greater focus on the specifics of quality instructional and assessment practices and how they work
- Technology-based learning environments & tools
 - Enable enhanced instructional & assessment practices
- Educating policymakers -- the need to focus on instruction and understand the limits and uses of tests

Assessment & Accountability Program Design: Making Success More Than A Miracle!

"I think you should be more explicit here in Step Two."

