



Role of Student Achievement Data in Public Engagement

Iris Weiss

Dawayne Whittington

for the

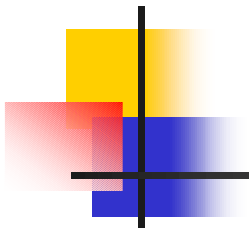
Special LSC Public Engagement Mathematics Conference

April 22, 2002

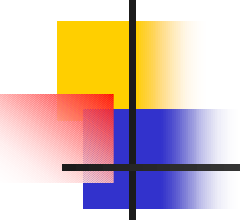
Las Vegas, Nevada



“Ideas matter in education.”



“Numbers matter in education.”



“The ideas that matter the most are the ones we choose to put numbers on.”

- Tony Bryk, University of Chicago



AEA Position

“High Stakes testing leads to under-serving or mis-serving all students, especially the most needy and vulnerable, thereby violating the principle of ‘do no harm.’ The American Evaluation Association opposes the use of tests as the sole or primary criterion for making decisions with serious negative consequences for students, educators, and schools. The AEA supports systems of assessment and accountability that help education.”



CRESST's Standards for Educational Accountability Systems

- Intended to evaluate existing systems and guide the design of improved procedures
- Represent compiled knowledge developed from several sources, including AERA, APA, and NCME
- Represent models of practice from 3 perspectives: research, practical, and ethical



CRESST's Standards for Educational Accountability Systems

Include 5 Categories

- Standards on System Components
- Testing Standards
- Stakes
- Public Reporting Formats
- Evaluation



CRESST's Standards for Educational Accountability Systems

Standards on System Components

- 1) Accountability expectations should be made public and understandable to all participants in the system.
- 2) Accountability systems should employ different types of data from multiple sources.
- 3) Accountability systems should include data elements that allow for interpretations of student, institution, and administrative performance.
- 4) Accountability systems should include the performance of all students, including subgroups that historically have been difficult to assess.



CRESST's Standards for Educational Accountability Systems

Standards on System Components

- 5) The weighting of elements in the system, different test content, and different information sources should be made explicit.

- 6) Rules for determining adequate progress of schools and individuals should be developed to avoid erroneous judgments attributable to fluctuations of the student population or errors in measurement.



CRESST's Standards for Educational Accountability Systems

Testing Standards

- 7) Decisions about individual students should not be made on the basis of a single test.
- 8) Multiple test forms should be used when there are repeated administrations of an assessment
- 9) The validity of measures that have been administered as part of an accountability system should be documented for the various purposes of the system.
- 10) If tests are to help improve system performance, there should be information provided to document that test results are modifiable by quality instruction and student effort.



CRESST's Standards for Educational Accountability Systems

Testing Standards

- 11) If test data are used as a basis of rewards or sanctions, evidence of technical quality of the measures and error rates associated with misclassification of individuals or institutions should be published.
- 12) Evidence of test validity for students with different language backgrounds should be made publicly available.
- 13) Evidence of test validity for children with disabilities should be made publicly available.
- 14) If tests are claimed to measure content and performance standards, analyses should document the relationship between the items and specific standards or sets of standards.



CRESST's Standards for Educational Accountability Systems

Stakes

- 15) Stakes in accountability systems (or incentives and sanctions) should apply to adults and students and be coordinated to support system goals.
- 16) Appeal procedures should be available to contest rewards and sanctions.
- 17) Stakes for results and their phase-in schedule should be made explicit at the outset of the implementation of the system.
- 18) Accountability systems should begin with broad, diffuse stakes and move to specific consequences for individuals and institutions as the system aligns.



CRESST's Standards for Educational Accountability Systems

Public Reporting Formats

- 19) System results should be made broadly available to the press, with sufficient time for reasonable analysis and with clear explanations of legitimate and potential illegitimate interpretations of results.

- 20) Reports to districts and schools should promote appropriate interpretations and use of results by including multiple indicators of performance, error estimates and performance by subgroup.



CRESST's Standards for Educational Accountability Systems

Evaluation

21) Longitudinal studies should be planned, implemented, and reported evaluating effects of the accountability program. Minimally, questions should determine the degree to which the system:

- builds capacity of staff;
- affects resource allocation;
- supports high-quality instruction;
- promotes student equity access to education;
- minimizes corruption;
- affects teacher quality, recruitment, and retention; and
- produces unanticipated outcomes.



CRESST's Standards for Educational Accountability Systems

Evaluation

22) The validity of test-based inferences should be subject to ongoing evaluation. In particular, evaluation should address:

- aggregate gains in performance over time; and
- impact on identifiable students and personnel groups.



Whether We Like It Or Not

- High-stakes testing is a reality.
- “No Child Left Behind” will result in more tests and higher stakes.
- The key is learning to work within this environment and building a strong case for what you believe matters most.



Research Credibility

- Flaws in a study can give opponents ammunition to discredit your entire effort.
- Credibility is key
 - Maintain objectivity, real and perceived
 - Ensure appropriate methodology
 - Acknowledge weaknesses in study design



The Public

- Engaging the public requires understanding their values and beliefs.
- Don't try to convince the public that your desired outcomes are the only important ones.



The Public

- Examine measures deemed important by both you and the public.
 - Problem-solving
 - Critical thinking
 - “Basic” skills



The Public

- The reform community has lost the rhetoric war – so choose your language carefully.
- You say:
 - “Authentic assessment”
 - “Conceptual understanding”
- They hear:
 - “Lack of rigor”
 - “We don’t care if kids can’t compute”



The Public

- When shown actual items and asked what types of questions they would like their child to be able to answer, most parents selected those requiring critical-thinking/problem-solving, not pure computational items.



The Public

- To win public support, you must show gains in measures the public cares about.
- Be careful about what you promise – set realistic expectations.



Methodological Credibility

- To be credible, your research must follow a scientific process
 - Research/evaluation needs to determine “if” and “how” programs work, not to show that they do.
 - Cherry-picking data is not research.



Methodological Credibility

- Need to design and implement the best studies possible
 - Build a chain of evidence that helps you link the outcomes with the reform interventions
 - Address rival hypotheses



Methodological Credibility

- While no single study will be perfect, multiple studies, each with different shortcomings, will help you build the case for success.



Ohio SSI

- Study 1 – controlled for volunteer effect
- Study 2 – controlled for school context
- Consistent results



Briars and Resnik

- Included a researcher who was not involved in the program.
- Included traditional test as well as New Standards.
- Built a chain of evidence, including comparing high and low implementers.
- Pointed out limitations in the study design.



In Summary...

- Policy makers are accountable for designing accountability systems that help improve teaching and learning.
- Education reformers are similarly accountable for providing information about the impact --and lack of impact-- of their reforms in order to help improve teaching and learning.