



On Test Quality...And Beyond

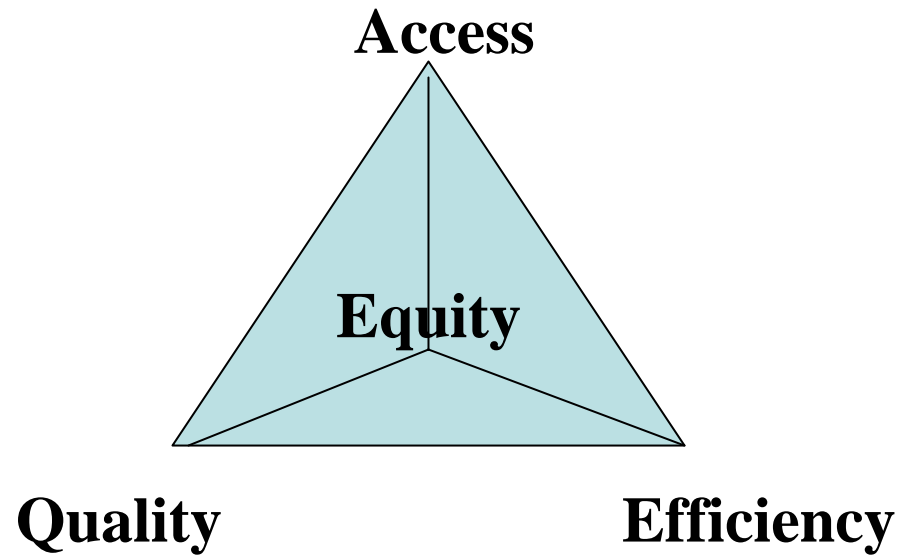
Henry Braun
Educational Testing Service

Presented at the

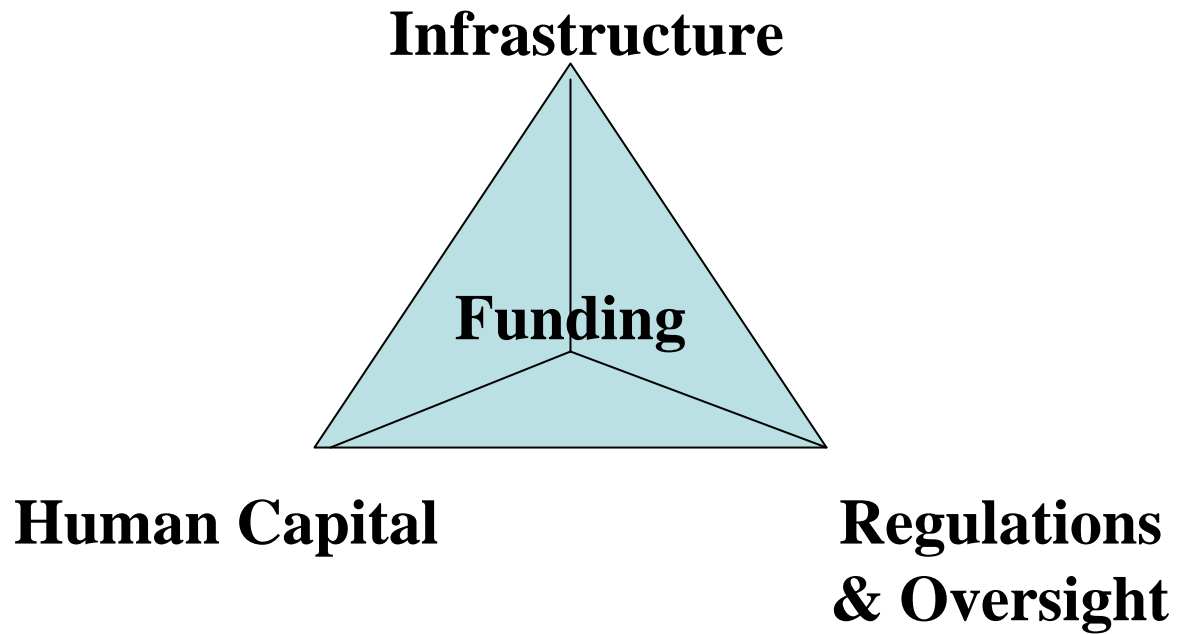
Carnegie Centennial

June 1, 2006

Education Policy - Goals



Education Policy - Means



The Roles of Testing

- Goals are set by society through legislative actions and bureaucratic interpretations.
- Testing is an instrument of education policy that usually operates through Regulations and Oversight.

Quality (1)

- Test results can be used to define the standards by which quality is determined (e.g., “Proficiency” is defined as obtaining a score of 80% or better on a test.)
- Test results can highlight where quality is lacking and indicate possible problems (e.g., students in a particular school are having trouble with algebra.)

Quality (2)

Test results are now being used for institutional and/or teacher accountability, based on the following logic:

“If good schools and good teachers are critical to student learning, then shouldn’t evidence of student learning – or its absence – tell us about the quality of schools and teachers?”

Models

Schools

Teachers

Students

Status	✓		✓
Growth	✓		
Value-added	✓	✓	

Only value-added models directly instantiate this logic by attempting to “isolate” the contributions of schools and/or teachers to student learning, as measured by patterns in test score trajectories.

There are methodological concerns related to interpreting the output of a value-added analysis as an accurate indicator of school or teacher effectiveness.

Essentially, they are related to the problem of making causal inferences from observational studies.

Test Quality (A Narrow View)

- Minimize
 - construct under-representation
 - construct irrelevant variance
- Improve
 - reliability
 - form-to-form comparability

Test Quality (A Broader View)

Consider the assessment system, which comprises (at least):

- curriculum & performance standards
- test specifications
- test items
- scoring
- reporting

Test Quality (A Yet Broader View)

- How does the current emphasis on reading and mathematics accountability impact school and teacher efforts?
- How does the standard-setting process affect results and reports?
- How does the use of “status models” influence resource allocation at each level of the system?
- What new issues arise with the introduction of a longitudinal component to accountability?

Campbell's Law

“The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures, and the more apt it will be to distort and corrupt the social processes it is intended to monitor.”

Poor test quality makes it more likely that the consequences of Campbell's Law will occur.

But poor assessment system design and/or implementation also invites resistance and corruption.

Systemic Validity (1)

Education policies are systemically valid if they result in decisions and actions that lead to progress toward one or more intended goals without causing regression with respect to other goals.

Systemic Validity (2)

Given a proposed accountability system, what is the “theory of action”? That is, how will the system accomplish its goals?
What is the empirical evidence to support this theory of action?
Is there empirical evidence to the contrary?

What are plausible alternative scenarios, incorporating possible negative unintended consequences? What are the probabilities and costs associated with these unintended consequences?

Systemic Validity (3)

Recognize the need for an interactive policy design process in which original design decisions are modified in light of analysis of alternative scenarios and anticipated costs (monetary and other).

Design modifications for an accountability system can include:

- Improve construct validity
- Broaden basis for evaluation
- Set realistic standards
- Enhance substance and appearance of fairness
- Allocation of resources
- Phase-in of policies while building capacity

Conclusion

Accountability for schools and teachers can play a constructive role in improving student learning, provided that reasonable steps are taken to enhance systemic validity.

The challenges are both technical and political and are made more daunting by ideological polarization in both educational research and policy-making.

The Question

How can The Carnegie Foundation, ETS, and other institutions with a deep commitment to educating all students help states and districts to meet these challenges?