Instructional Sensitivity and Accountability

Marianne Perie, CETE
November 15, 2013
Claim 1. The content standards have been designed to facilitate learning in a subject matter by building on prior knowledge through known subject and cognitive development.

Claim 2. The assessment has been designed to yield scores that reflect students' knowledge and skills in relation to academic standards.

Claim 3. The assessment is administered and scored as intended.

Claim 4. The assessment performance levels reflect meaningful differences in students' proficiency.

Claim 5. Assessment scores and performance levels accurately reflect the level of knowledge and skills each student has.

Claim 6. Change in individual assessment scores from one year to the next reflects the level of learning during the school year.

Claim 7. Assessment scores can be used to evaluate student proficiency in the subject matter.

Claim 8. Changes in assessment scores can be used to evaluate teacher effectiveness.

Claim 9. Distribution across proficiency levels and changes in that distribution can be used to evaluate school and district quality.

Students leave the grade/school having gained sufficient knowledge and skills required for the next phase of education or training.

Teachers are recognized for their effectiveness and given feedback to improve.

Resources are appropriately allocated to schools and districts to maintain or increase their quality.
Goal 1

- Students leave the grade/school having gained sufficient knowledge and skills required for the next phase of education or training.
- Always a goal - not related to federal accountability efforts.
- Not as dependent on instructional sensitivity - question is...do they know the material sufficiently or not?
Goal 2

- Teachers are recognized for the effectiveness and given feedback to improve.
- Became a goal under RTTT and a requirement of receiving a waiver.
- Key is attribution:
  - Can student scores be linked to instruction?
  - Can instruction be linked to an individual?
- We’ll come back to this...
Goal 3

- **Resources are appropriately allocated to schools and districts to maintain or increase their quality.**
- Major outcome under NCLB and the waivers.
- Dependent on instructional sensitivity to a lesser degree than Goal 2.
- Takes in other factors to learning such as school climate and resources.
Teacher effectiveness

Assumption 1. Tests produce valid scores reflecting what students know and can do.

Assumption 2. Tests were designed to be sensitive to instruction.

Sub-Assumption 2.1. Items are written to reflect real learning and not just generalized knowledge.

Sub-Assumption 2.2. Items measure new knowledge and efficiency of skills.

Assumption 3. Change in test scores can be attributed to new learning.

Assumption 4. Test design allows for interpretation at small enough grain size to provide effective feedback.

Assumption 5. Changes in test scores can be attributed to a particular teacher.

Teachers are recognized for their effectiveness and given feedback to improve.
Concluding Thoughts

- Tests that are sensitive to instruction can be more validly used to measure changes in student learning.
  - *As long as instruction is focused on the big ideas in the standards and not trivia.*

- Regardless of the assessment, it’s the attribution that gets us in trouble.
  - Can we be sure that learning is attributed to school activities?
  - Even if so, how do we attribute that learning to a particular teacher?
More concluding thoughts

- Is the ease with which instructionally sensitive assessments can be attributed to a particular teacher dependent on the subject and grade level?
  - Elementary grades, one teacher model
  - Secondary grades, complexity of subject
    - Subjects such as calculus are probably only taught by the calculus teacher
    - Subjects such as “analytic writing” are probably taught by multiple teachers
Education is what remains after one has forgotten what one has learned in school.

-Albert Einstein