Educational Testing and Assessment: Context, Issues, and Trends
The role of testing and assessment in Education

- Testing and Assessment are often used as measures of quality of learning, teaching, and education. Some purposes include:
  - Yardstick of individual progress in classrooms
  - Placement/filter of educational opportunity
  - Barometer of national educational condition
  - Evaluation of quality of teaching
  - Diagnosis of learning problems
  - Agents of school reform
Testing/Assessment as agents of school reform

- Educational Policymakers are very assessment-focused. Why?
  - An inexpensive way of bringing about school reform
  - Testing can be externally mandated and enforced (compared to changing standards of teaching)
  - Testing/assessment reforms can be quickly implemented
  - Results are visible and amenable to press reports
In what ways have test-based reforms been implemented in last 50 years?

• Though testing has been around and popular for a long time, let us specifically examine the last 5 decades

1. Title I – aimed to improve the quality of education for, and educational achievement of students who are at-risk (migratory, poor, disabilities, behind in reading, Limited English Proficient). There is a strong focus on accountability of schools – major push for testing – both in Fall and Spring semesters
In what ways have test-based reforms been implemented in last 50 years? (continued)

2. Minimum-Competency Testing. Arising from concerns that high school graduates do not have the basic skills, MCT was implemented in the 1970s. The idea is to hold students back (cannot graduate) unless they pass these tests. We have now moved beyond MCT – leading to tests and assessment that measure higher levels of learning.
In what ways have test-based reforms been implemented in last 50 years? (continued)

3. “A Nation At Risk” published by National Commission on Excellence in Education (1983) – emphasized the role of tests in documenting student learning and holding schools responsible/accountable. Resulted in school and district “report cards”. This high-stakes testing, resulted in practices like teaching the curriculum of the test, teaching test-taking skills, and using the format of externally mandated tests in home-made tests. Everyone became above average! – Lake Wobegon Effect (see page 6).
In what ways have test-based reforms been implemented in last 50 years? (continued)

4. Standards-Based Reform (SBR). Setting up world-based standards, this movement (1990s) was instrumental in further emphasizing assessment, but at a higher level of response /thinking – e.g., performance-based assessment. Let’s examine the four aspects of SBR:

- Content and Performance Standards – content standards specify what should be learned at each grade; performance standards state what students should be able to do (e.g., TEKS is one type of state response to this)
In what ways have test-based reforms been implemented in last 50 years? (continued)

4. Standards-Based Reform (SBR).

- Performance-Based Assessment. A result of the push for alternative forms of assessment (AKA authentic, direct assessment). The idea is to make assessment something that is worth teaching to (i.e., teaching to multiple-choice tests is not acceptable practice. Since schools usually end up teaching to the assessments, make these assessments higher level, more authentic).
In what ways have test-based reforms been implemented in last 50 years? (continued)

4. Standards-Based Reform (SBR).

- High-Stakes Accountability Mechanism. SBR brought about higher emphasis on high-stakes accountability for world-class standards (as opposed to minimum competency testing) – and usually show up as: (a) ratings of schools; (b) school and teacher bonuses; (c) higher standards for grade-to-grade promotion and graduation.
In what ways have test-based reforms been implemented in last 50 years? (continued)

4. Standards-Based Reform (SBR).

- Inclusion of All Students. Usually, children with disability, Limited English Proficient, and other at-risk children were excluded from testing. With SBR, the focus became to include all students in this testing/accountability initiative. In fact, NCLB indicates that schools with less than 95% of students being tested will be automatically placed in the “needs improvement” category. However, accommodations for students with disabilities are allowed
5. No Child Left Behind (NCLB). Seemed to extend the Title I requirements to all students/schools. This requires states to identify content standards, demonstrate that they are challenging, and to identify performance standards at three levels – basic, proficient, advanced. The goal is to ensure that all students are “proficient” by 2014. Each year schools are rated to determine their progress toward this goal (Adequate Yearly Progress). See handout on NCLB.

In what ways have test-based reforms been implemented in last 50 years? (continued)

5. No Child Left Behind (NCLB) – Key Features
   a. challenging state standards in reading and math
   b. annual testing for all students grades 3-8
   c. annual statewide progress objectives ensuring that all groups reach proficiency in 12 years (by 2014)
      -poverty, ethnicity, race, disability, and LEP (Limited English Proficient) groups
   d. schools that do not make adequate yearly progress toward statewide proficiency goals will be subject to improvement, corrective action, and restructuring (over 3 years).
In what ways have test-based reforms been implemented in last 50 years? (continued)

5. No Child Left Behind (NCLB) – Key Features

e. More choices for parents – schools which persistently fail must provide supplemental services through other public agencies; schools identified for improvement must give parents a choice to attend a better public school – the school must provide transportation; schools that persistently fail (5 years) run the risk of reconstitution

f. States have flexibility in setting standards, implementing major programs (which may be funded by NCLB/USDOE)
In what ways have test-based reforms been implemented in last 50 years? (continued)

6. Race To The Top funds – the USDOE has created a bigger mandate for adoption of data-driven instruction and decision-making by requiring these elements to be included in any educational reform proposals that are funded by the Race To The Top funds. This includes the requirement of evaluating teachers based on improvement in student achievement scores. In NY, school districts have developed formulae for the APPR, a percentage of which is based on improvements in student standardized test scores.
Some other trends in assessment

• Technological advances in testing and assessment
  – Computer-based testing is getting highly popular because of its efficiency, adaptive testing capabilities, quick feedback. In addition, computers can be used to administer test items with simulations. The problem is that many schools still do not have enough computers. Also, if computer-based tests are taken without any practice, scores do go down. This can be a very powerful tool in testing using authentic cases.
Some other trends in assessment (continued)

- Public concern about testing and assessment. Due to local control over testing programs, parents often play a role in a school district’s testing programs. These concerns can be presented in four points:

1. Nature and Quality of Tests
   a. reading too much into multiple-choice questions, artificiality
   b. Tests measure only limited aspects
   c. Test scores may be subject to “error” (fatigue, illness, test-wiseness, familiarity)
Some other trends in assessment (continued)

2. Effects of Testing on Students
   a. Test anxiety (what should teachers do? – practice, adequate preparation, liberal time limits)
   b. Tests categorize and label students. Pigeon-holing students may create self-fulfilling prophecies. Teachers may also create and communicate expectations to the negatively labeled child. Treat test scores as measured of limited aspects of student performance.
   c. Tests damage students’ self-concepts. Focus on effort to explain failures. Do not over-generalize from scores.
   d. Tests create self-fulfilling prophecies.
Some other trends in assessment (continued)

3. Fairness of tests to minorities.
   a. Absence of bias (does the test equally predict performance / success for all groups?)
      i. Absence of content bias
      ii. Appropriate use of information (i.e., use test data to improve learning, not to reward or punish groups of students)
   b. Procedural fairness (do the test procedures allow everyone to show what they know?)

Two other ways in which bias is defined are:

   c. Are the test results based on an equal opportunity to learn?
   d. Are results equal across all groups?

Note that the use of test results is more important.
4. Gender Fairness

There is a difference in average scores of males and females on tests – e.g., PSAT (used to identify national merit scholars).

Does this mean that the test is biased? OR do these differences show inequities in opportunities to learn? Females take fewer math scores than males do.

For example, if we use upper body strength as a test of qualification to be a truck driver, which group is more likely to be selected? Does that mean that the test of upper body strength is biased or that the use of the test results is biased?