

Student Academic Testing

Dr. Joseph Graves, Secretary of Education

Why does it matter?

- **Testing is a constant topic of discussion in schools**
 - New importance under No Child Left Behind and subsequent iterations of the ESEA
 - As federal involvement education lessens or morphs, including the US DOE innovative statewide assessment program, the topic is of increased interest
 - Currently, it is also of increased interest as data from state testing has impacts on various education reforms
- **The following is a mix of background on academic testing and relevant information to topics under discussion.**

Background: Five Types of Testing

1. Norm-Referenced
2. Criterion Referenced
3. District Administered
4. Teacher Developed/
Classroom Assessments
5. Aptitude Testing

Norm-Referenced Testing

- **Definition: An assessment created for the purpose of comparing an individual test-taker to his educational peers**
 - Iowa Tests of Basic Skills, Iowa Tests of Educational Development, ACT, SAT, etc.
 - Scores from 1-99, based on percent of students the test-taker is doing better than. Scores assigned across the entire range.
- **Tests have been field-tested and “normed;” also have gone through statistical testing to demonstrate validity and reliability**
 - **Validity:** The extent to which an assessment tests what it purports to test
 - **Reliability:** The extent to which an assessment taken by the same test-taker will result in the same score.

Criterion-Referenced Testing

- **Definition: An assessment created for the purpose of measuring mastery of specific skills, concepts.**
 - Smarter Balanced Assessment, Dakota Step
 - Theoretically, scores could be 0-100. Every student could be at 100 if they all mastered the skills.
- **These tests have been field-tested and “normed;” they also have gone through statistical testing to demonstrate validity and reliability**
 - Validity: The extent to which an assessment tests what it purports to test
 - Reliability: The extent to which an assessment taken by the same test-taker will result in the same score.

District-Administered Testing

- **Definition: Administered at the district level multiple (usually 3) times per year, to measure student progress in a school year on a specific topic**
 - DIBELS (Dynamic Indicators of Basic Literacy Skills), aka Acadience Reading; and others
 - Often involves individual testing with a school employee, testing the student's master of one or more of the 5 SOR foci
- **Achieves Face Validity, Can be Susceptible to Reliability Issues**

Teacher-Developed/ Classroom Assessments

- **Definition: quizzes and various tests teachers use (and often write themselves) to determine student grades or inform subsequent instructional plans**
 - Comprises most of the test's students take in school – the end-of-chapter exam in science, the weekly spelling test, etc.
 - Scores are based upon the individual teacher's scale.

Aptitude Testing

- **Definition: Assessments designed to highlight individual student abilities, proclivities, interests**
 - ASVAB, MAPP, etc.
 - The tests result in scores but the results are more about whether you are good at certain skills, in certain environments, in certain interest areas.
 - Irrelevant to today's discussion

Why is Accountability and Academic Testing of Interest Now?

- **Scores on Smarter Balance Assessments Help Form Public Perception**
 - 51% ELA Proficiency: “Half our students can’t read.”
 - 44% Math Proficiency: “Our students won’t be able to compete in the global marketplace.”
- **Ambiguity about Federal Involvement**
 - Will the Feds return accountability to the states?
 - Will the Federal testing requirement go away or be revised?

Possible Solutions under Smarter Balanced Assessments

- **Mark “proficiency” at Levels 2 through 4 rather than Levels 3 & 4.**
 - Current Levels: 1 – Below Basic, 2 – Basic, 3 – Proficient, 4 – Advanced
 - Current: 51% proficiency, ELA, 44% proficiency, math
 - New: 75% proficiency, ELA, 71% proficiency, math
 - New number is “better” but it presents new numbers with no actual change in student achievement.
 - Looks fishy.
 - Depending upon how you use it, it *is* fishy
 - Possible Solution: publish it both ways

Possible Solutions

Alter Cut Scores

- What level a student's score place him into is based on "cut" scores: a number of correct responses on the test determined by a committee of educators in the past.
- Moving cut score requirements up or down would change the percentage of students who qualify for specific levels.
- Same concern: You can make the argument that the new cut scores are more representative of reality, but it reflects no actual change in student achievement.

(fishy)

Clarify the meaning of the 4 levels through public communications efforts

- "Basic" cut off is too blunt of an instrument: many level 2 students are on their way to proficiency.
- Strategy could be a forthright means of correcting misperceptions
- But the "number" and level names are powerful, hard to overcome

Possible Solutions

Publish NAEP (National Assessment of Educational Progress) with State Assessment Data

- NAEP – also called *the Nation's Report Card*
- Publishing NAEP Scores with Smarter Balanced scores provide frame of reference
- South Dakota National (NAEP) Standing:

2022	ELA Rank	Math Rank
4 th Grade	14 th	13 th
8 th Grade	12 th	4 th

2024	ELA Rank	Math Rank
4 th Grade	31 st	11 th
8 th Grade	16 th	6 th

Ok, That's Great,
but I Still Have
Questions.

Can a school district include, in their reports, information beyond the state accountability data?

Yes. While all schools must communicate accountability data to their parents and the public, they are free to include other metrics of their choosing.

Why would testing data from one kind of test not yield data consistent with data from another test?

This can and does happen frequently.

- A student with lots of progress on DIBELS during a current school year might not see the same kind of growth on Smarter Balanced from year to year.
- This could be because:
 - The two are measuring somewhat different things;
 - They are testing a learning period different from one another; or
 - There are validity and/or reliability issues.

Are South Dakota students doing well in reading and math?

That depends upon the angle from which you are looking at them.

- South Dakota students tend to outscore students nationally.
- The older they get, the more they outpace them.
- Yes, but . . . 44% of students proficient in math by the rigorously set cut scores is not much to brag about.

Why do parents
sometimes seem
unconcerned about
the achievement of
their child's school
even when
proficiency rates
seem low?

Experienced educators in the field suggest that parents are far and away most concerned about how their child is doing.

When they feel their child's scores or levels are accurate or pleasing, they are receptive to the testing.

Just *One More* Question:

What is *the* solution to this issue of student academic testing?

Opinions will vary, of course, but one thing would resolve all concerns: a better understanding by all interested parties of just what test scores mean.

- That requires us to better communicate with stakeholders: parents, communities, the media, etc.
- If we can do that, if we can reach an agreement on just what the data means, these issues should disappear.
- Better yet, they should point us in a mutually agreed upon decision on how best to *improve* student achievement.