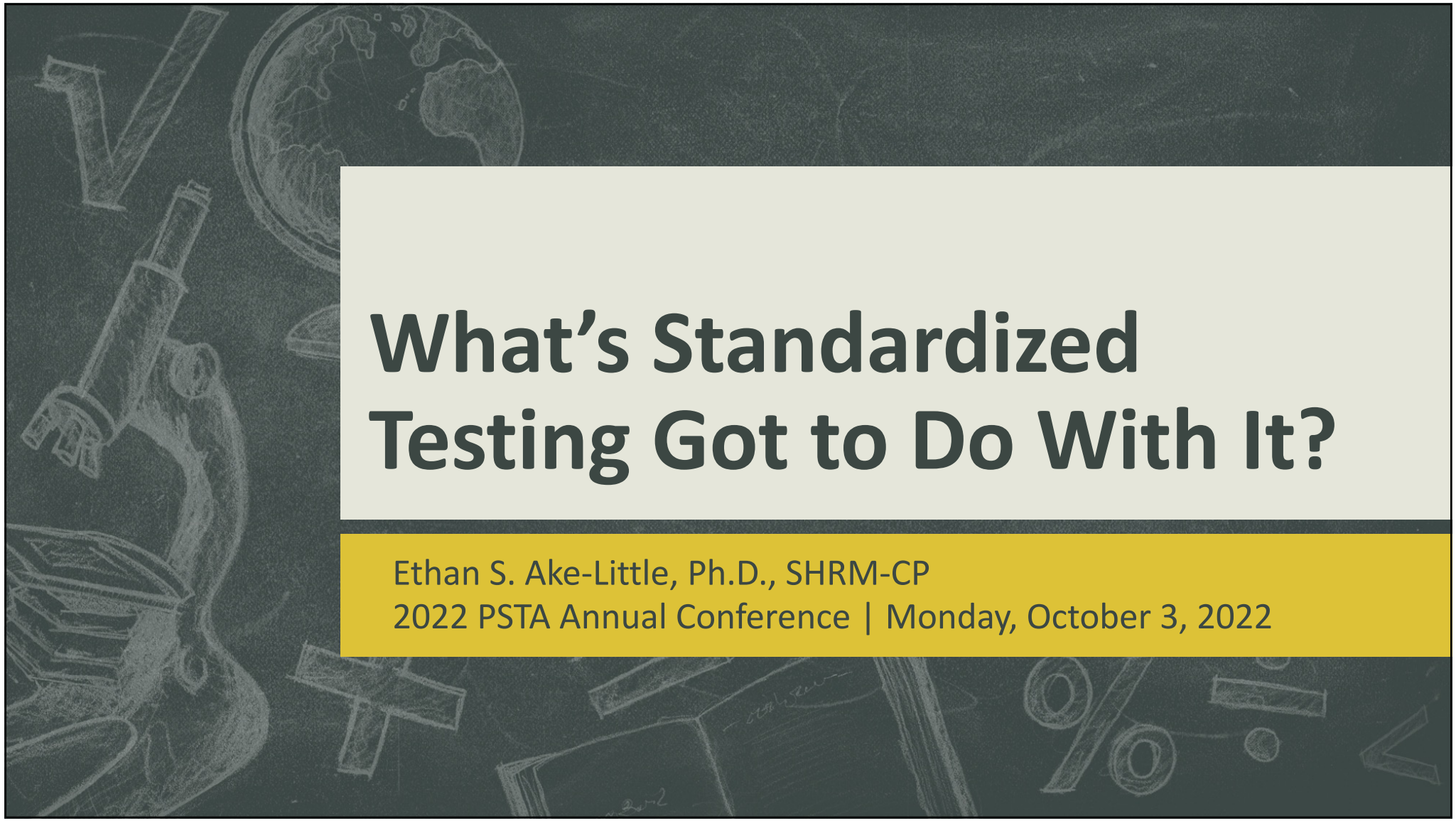


What's Standardized Testing Got to Do With It?

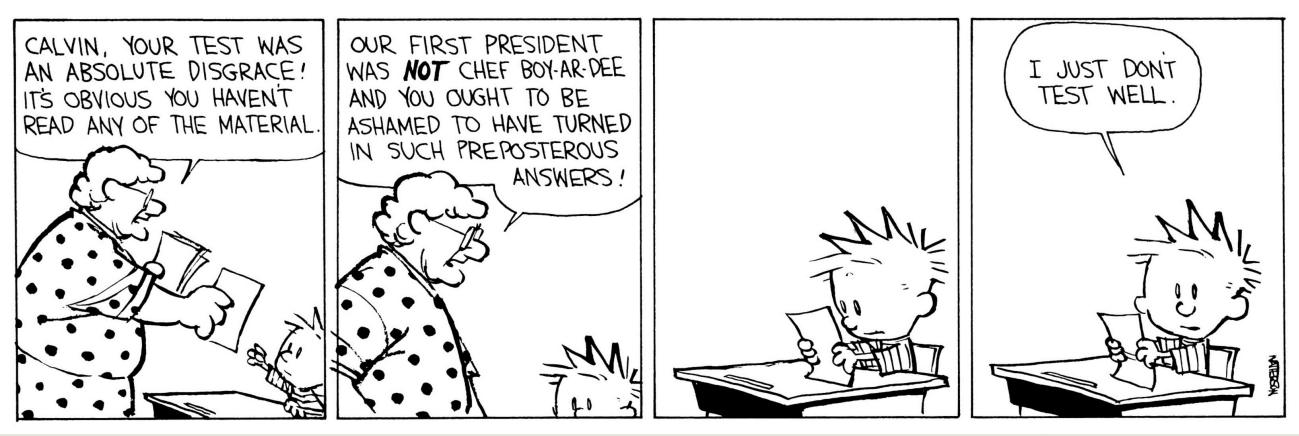
Ethan S. Ake-Little, Ph.D., SHRM-CP

2022 PSTA Annual Conference | Monday, October 3, 2022

Presentation Overview

Objectives

- ✓ Understand the history and logic underlying standardized testing creation and development
- ✓ How the relationship between curriculum and instruction as measured by standardized testing is a derivation of the logic used in IQ testing
- ✓ Discuss ways in which you (as school administrators or teacher leaders) can improve teaching and learning to better prepare students for standardized assessments



Agenda

- ✓ The Logic of Standardized Testing (15mins.)
- ✓ The Relationship Between Curriculum & Instruction and Standardized Testing (15mins.)
- ✓ Practical Advice and Q&A (10mins.)

The Origins of Standardized Testing

Standardized testing has its roots in IQ testing.

- An IQ test is a ratio of a person's mental age to their chronological age (multiplied by 100)
- There were many attempts at measuring intelligence in the 19th century, but it was the work of Alfred Binet (University of Paris) and Lewis Terman (Stanford University) who created the progenitor of modern IQ testing – the Stanford-Binet Test.
- Unfortunately, what might have started as a scientific endeavor became a vehicle for racism and exclusion. (e.g., Moron = 40-55; Retarded = 55-70; Talented = 115-130; Genius = 130-145).

Translating IQ Tests to Subject Matter Exams

Standardized tests in schools are modified IQ tests.

- The Stanford-Binet Test has been revised/ updated over the decades but all IQ tests, including standardized tests, are based on the ratio of mental age to chronological age; in the context of schooling, that means the ratio of "mastery of content knowledge" to "grade level expectation".

- The part of that ratio that is "grade level expectation" is better known to us as curricular standards (e.g., PSSA Grade 8 standards). The assumption here is a **PROFICIENT** is a 1:1 ratio (you, the student, have mastered what you are expected to know at your grade level in this subject area); **ADV PROF** is >1:1 ratio whereas **BASIC/ BELOW BASIC** is <1:1 ratio.

CONTENT AREA	BELOW BASIC	BASIC	PROFICIENT	ADVANCED
Algebra 1	1200-1438	1439-1499	1500-1545	1546-1800
Biology	1200-1459	1460-1499	1500-1548	1549-1800
Literature	1200-1443	1444-1499	1500-1583	1584-1800

- All standardized tests are "normed"; the scale always stays the same *but* it is the content knowledge that changes.

- Someone who scored **PROF** on the Keystone Biology in 2012 might not have the same knowledge base as someone who scored the same in 2045.

You(r Students) vs. the World (of Students)

For standardized test scores to be accurate, they need to have a large sample size.

- We know that the *individual* score is a ratio between "mastery of content" versus "grade level expectation" but that score is only considered to be accurate if that test has been proven legitimate.
- As in any scientific experiment, the "experimental" group is your students; standardized tests often include questions that do *not* count towards the score but may be refined for future test use.
- The more individuals tested, the bigger the sample size. The bigger the sample size, the more accurate the test.

Who is the reference population?

- ✓ *Pennsylvania:* PSSA Math/PSSA Science; Keystone Algebra I/Biology
- ✓ *National:* P/SAT Math, ACT Math/Science
- ✓ *International:* IB or AP Math/Science

Standardized testing "works" in content heavy subjects (e.g., chemistry. American History, Algebra II) since there is a common curriculum (knowledge base).

Their value starts to decline in creative areas (e.g., painting, theater, music) since the creative act is difficult to standardize.

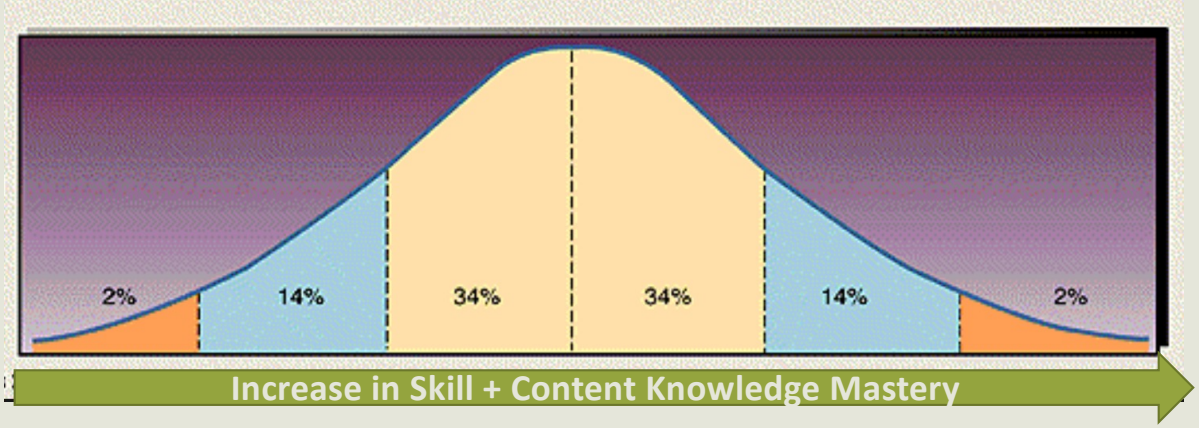
The Denominator: Grade Level Expectation

(Maybe) Good News! You don't have to worry about what is considered "grade level" appropriate in your subject area – the state has already done this for you!

- We know these grade level expectations as "standards" and the Pennsylvania Department of Education provides them via the Standards Aligned Systems (SAS) website.
- These standards are the basis of what we call the curriculum; which is the set of all lessons, teaching materials, assessments, etc. which (in theory) originate from the standards.

What happens if I can't cover all the standards (finish the curriculum)?

- Since the denominator is set (by the state), that means your students will not be able to master this curriculum.
- So, the denominator stays the same and the numerator *decreases*, resulting in a lower score (less proficiency). ☹



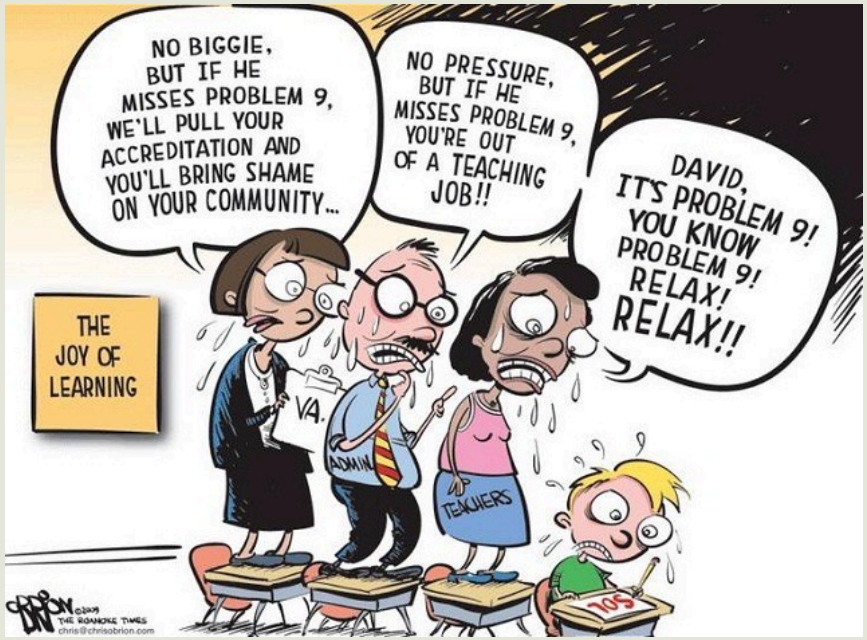
The Numerator: Mastery of Standards

Here's the tricky part: The state believes that you are responsible for a students' ability to understand (master) the standards.

- Because of this logic, you (as the educator) must do two (2) things at once: teach the standards (and, by extension, deliver the curriculum) as best possible *and* be responsible for mastery of the material.
- In fact, this belief that *you* are responsible for student mastery is the basis for another statistical measure called *value-added* (i.e., your teaching should lead to an increase in student achievement longitudinally).

$$\frac{\text{INSTRUCTION}}{\text{CURRICULUM}} = \textit{PROF}$$

Unfortunately, this assumption that *you* are responsible for a student's ability to master the standards has led to some misguided policies.



So, Some (Hopefully) Helpful Advice…

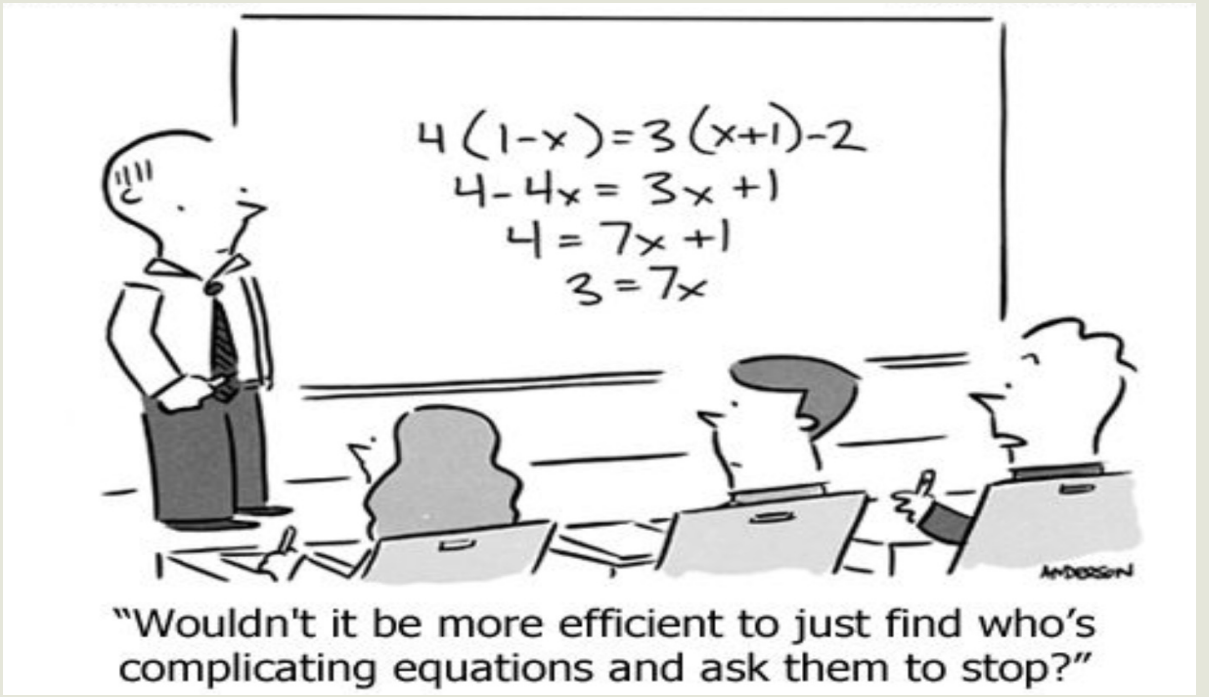
Data Driven Instruction

- ✓ Attempting to improve instruction without attention to assessment is futile!
- ✓ Internal assessment need to be (1) mapped in conjunction with curriculum to ensure tight alignment between the two and (2) should be in a format that also aligns with major external assessments.
- ✓ Data driven instruction can face resistance from teachers, but you don't need buy in; the results are the buy-in.

DO:

- ❖ Ask teachers teaching common subject areas to create unified assessments so that all students have a common benchmark.
- ❖ Encourage/make available data analysis related professional development for administrators and teachers.
- ❖ Provide teachers time (particularly on professional development days) to work within their departments or grades to analyze assessment results and modify instruction.

Examining assessment results at the end of the year is like attempting to diagnose a patient by performing an autopsy on a dead individual – **the damage is done and there isn't a second chance for the student.**



Ethan S. Ake-Little, Ph.D., SHRM-CP
Director of Human Resources | Wallingford-Swarthmore SD
State Representative | PSTA
Email: ethanake@temple.edu