Validating Indicators of College Preparedness: Ready or Not?

Presentation to:
THE 2008 REIDY INTERACTIVE LECTURE SERIES:
Validating Assessment and Accountability Programs
Thursday, September 25, 2008
Portsmouth, NH

Presenter:
Laress L. Wise
Validating Content and Performance Standards

• NCLB has increased attention to validity issues:
  – The validity of assessments as indicators of student mastery of targeted content
  – The validity of accountability systems as measures of school performance

• But little or no attention is being paid to the validity of the content and performance standards that underlie our assessment and accountability systems
  – Inference to be validated is that the content standards are important for students to master
  – Rarely is the rationale for these standards more sophisticated than “Because we said do”
Improving the Rationale for Content and Performance Standards

• Decisions on the content and level of achievement that define proficiency are not really *criterion referenced*
  – May involve business community as well as teachers and curriculum experts, but they are rarely asked to provide a rationale, let alone any evidence for proposed requirements
  – Some states align standards across grades, but without any empirical evidence of the relationship of mastery at one grade to readiness for content at the next grade
    • Lack of vertical alignment also complicates measuring growth
  – Current NAEP efforts to define “readiness” for 12th grade students are moving slowly and are not on the radar of most states
    • Framing student achievement expectations around readiness for college, work, and the rest of our lives
Rationale for Content and Performance Standards (Continued)

• Decisions on the content and level of achievement that define proficiency are not really normative either
  – Limited consideration of what other states require
  – Almost no consideration of what other countries require

• Without a rationale for prioritizing areas of knowledge and skill there is a tendency to just throw everything in
  – Resulting in mile-wide, inch-deep standards and curriculum
  – Making it difficult to assess the entire domain with a modest length test

• Are we teaching and measuring the right things?
Agreement on What to Teach Has to be a Significant Step Forward

You've got to be very careful if you don't know where you're going, because you might not get there.
Current Validity Evidence for Content Standards

• Peer review requirements:
  – Who makes the judgments about required content
    • Sufficient involvement of stakeholders
    • Diversity/Representativeness of panel members
    • Broad review
  – Adoption by appropriate policy-makers
An Alternative Model for Validating Content Standards

- Start with what we want students to know and be able to do after high school
  - Link to important post-high school outcomes
    - Success in earning a college degree
    - Career success: training and advancement in fulfilling careers
    - Citizenship
    - Success is leisure and avocational pursuits.

- Link objectives for earlier grades to success in achieving targeted knowledge and skills by the end of high school
  - Backwards mapping
  - Combined with forward-mapping of what we think students are able to learn at different ages (and grades)
College Preparedness Indicators

• College Preparedness is Important!
  – Kirst research shows prevalence and cost of need for remediation in college.

• Indicators that many students are NOT prepared suggest a failure of our K-12 education system.
  – Increased concern about international competitiveness (global economy)
  – Increased influence of the business community
Predictive Validity Evidence vs. Content Validity Evidence

• Predictive Validity Model – Employment Tests
  – When selection tests have adverse impact for protected groups, employers must demonstrate a “business necessity” for test use by showing a significant correlation between test scores and success in training or performance on the job. That’s it!
    • Inference: Higher scores mean better job performance.

• Content Validity Model – State Assessments
  – States perform alignment studies demonstrating coverage of targeted content standards.
    • Inference: Higher scores mean greater mastery of targeted knowledge and skills
Validity Model for Readiness Measures

• What’s the inference to be validated?
  A. Higher scores mean greater likelihood of gaining entry to college?
  B. Higher scores mean less need for remediation before taking credit-bearing courses.
  C. Higher scores mean mastery of specific knowledge and skills required for success in credit-bearing courses.
     – Hint: Always go for option C

• Correlational versus Causal Models
  – Correlation is not enough if we want to imply that school efforts to improve individual readiness, as measured by the test, will lead to increased ability to succeed in credit-bearing courses.
  – Example: A 1950’s study showed a significant correlation between the viscosity of asphalt and the incidence of polio.
    • But the real link was higher summer temperatures leading to increased use of community swimming pools.
    • Increasing the thickness of asphalt would not change the temperature and so would not decrease the use of pools and the incidence of polio.
Specific Knowledge and Skills versus “Ability to Learn”

• Recent changes in the SAT move toward assessing skills that are taught (responding to Atkinson/UC System)
  – And away from more general abilities that might indicate ability to learn.
  – Not can you learn, but have you learned (what was taught).
  – Reinforced by recent NACAP Report (http://www.nacacnet.org)

• Not without some controversy
  – Learning ability measures could provide a way of identifying possible success for students with more limited opportunities to learn.

• But remedial courses are about correcting specific knowledge and skill deficits, not about improving more general learning abilities.
  – Constructing an argument and writing an essay laying out a position
  – Or basic mathematical knowledge (e.g., exponents) necessary to learn calculus
Validity Model for College Readiness Measures

1. Start with a rationale for the required knowledge and skills
   - Likely basic verbal and quantitative skills

2. Link to K-12 curriculum
   - And vertically align K-12 curriculum to lead up to required knowledge and skills

3. Link mastery of these skills to decisions regarding need for remediation (clearer than admissions)
   - Could be content judgments, but correlations based on empirical data might be stronger

• Not the **only** goal of a K-12 education
  - Need to value science, social studies, foreign language, citizenship, etc.
Summary

• Validation of Content Standards should go well beyond a review of the adoption process

• Ensuring that students are ready for college work (without remediation) is a reasonable goal (among others) for K-12 education
  – Vertical articulation could provide a rationale for goals at earlier grades building to college readiness by 12th grade
  – Readiness is more than just high SAT/ACT scores!

• Validating college readiness standards could require both
  – Content judgments about prerequisite skills for credit-bearing college courses
  – Empirical evidence linking mastery at each grade (through 12) to success at what comes next
Teaser for Friday Morning Session: Example #3: NAEP Readiness Validation

- Math and Reading Content Specifications reviewed by Achieve and others with regard to preparedness for:
  - College (without having to take remedial courses)
  - Employment Training: for challenging careers
  - Military Service

- Technical Panel convened to recommend studies to validate the resulting measures.
  - Chaired by Michael Kirst and including both educational policy, measurement, and industrial psychology experts.
  - Report due at the November 2008 meeting of the National Assessment Governing Board (NAGB)