

VERTICALLY-ARTICULATED GRADE LEVEL EXPECTATIONS AND TEST SPECIFICATIONS

Stanley Rabinowitz, Ph.D.

WestEd

srabino@wested.org

**Presentation at the 2004 Edward F. Reidy, Jr., Interactive Lecture Series
October 7-8, 2004, Nashua, NH**

PRESENTATION SET OF QUESTIONS

- 1. What concerns about vertical scales need to be addressed for full implementation and interpretation?**
- 2. How do various assumptions affect the confluence of vertical scales and value-added models?**
- 3. How can thoughtful test specifications positively affect the concerns in questions 1 and 2 and assessment and accountability programs overall?**

2004 RILS EXPLORATION

If the goal of the assessment/accountability program is to measure whether:

- students have achieved a year's growth within an academic year (or other interval—e.g., by graduation); *and/or*
- schools have provided value-added instruction; *then*

Vertical Scales and Value-added Models have great intuitive appeal, *BUT*

PROBLEMATIC REQUIREMENTS OF VERTICAL SCALES

- ***Construct***, in particular construct invariance and unidimensionality across the *full* range of grades within a content area
- ***Alignment***, structure of content standards across grades
- ***Assessments***, consistency across grades (“mixed models”)?
- ***Design***, especially testing students on content they have not been taught or taught in previous years

Role of test specifications in overcoming problems

CHALLENGES OF A VALUE-ADDED SYSTEM: *VALUES*

- **Concept of “one-year’s growth” (across grades, content areas, students)**
- **Individual student growth most important measure of school effectiveness (student by student)**
- **Classroom/teacher effect most important factor in student success (vs. school, district, home, community)**
- **Student test scores are a proxy for effective teaching**
- **Changes in standardized test scores are what is valued**
- **Additional challenges: assessment, data systems, political**

TEST SPECIFICATION CONSIDERATIONS

- **Start with clear understanding of what content is important and what is less so (vs. unimportant)**
- ***Specificity* of specifications**
- ***Articulation* of specifications (not just standards)**
- ***Performance levels* need to be considered**
- **All *purposes* of assessment/accountability program need to be taken into account**
- **Effects on *special populations* (SE, EL)—
interaction with assessment development
(*Universal Design*)**

TEST SPECIFICATION CONSIDERATIONS

(cont.)

- **Development of assessment blueprint:**
 - **three dimensional matrix**
 - **cognitive complexity**
- **Alignment studies (formal and ongoing):**
 - **vertical and horizontal alignment of standards**
 - **assessments to blueprint**
- **What is the meaning of a test score?**
- **How will test results be interpreted?**