Challenges in Assessing Student, School, District, State, and National Performance at the High School Level

Lauress Wise

Human Resources Research Organization (HumRRO)

Edward F. Reidy, Jr. Interactive Lecture Series

Nashua, New Hampshire

September 29, 2005

High School Assessments

Presentation outline

➢ **High School Assessment Questionnaire**

  ▪ **Self Assessment**

➢ **Key Technical Issues**

  ▪ **Focus on Issues Unique to High School Assessments**

➢ **Some Alternative Models**

  ▪ **Table Talk**

High School Assessment Questionnaire

➢ **Please take a few minutes to complete my high school assessment questionnaire.**

  ▪ What questions are you interested in answering?
  ▪ What types of assessments are you interested in using at the high school level?

➢ Intended to help you frame questions to ask during this presentation.

  ▪ *(This is the interactive part).*

Questions in Search of Data

- ➢ About Students:
  - ■ Mastery of basic skills
  - ■ Mastery of advanced skills
  - ■ Diagnosis of strengths and weaknesses
  - ■ Readiness
  - ■ Non-academic traits
- ➢ For:
  - ■ Promotion/graduation decisions
  - ■ Targeting remediation
  - ■ Improving instruction

- ➢ About Schools (states, the nation, the world):
  - ■ Average score of students
  - ■ Value-added
    - ✓ All students
    - ✓ Targeted groups
  - ■ Who is being left behind?
- ➢ For:
  - ■ Rewarding high-performing schools
  - ■ Shaming other schools (or worse)
  - ■ Teacher pay decisions
  - ■ Improving instruction

The Many Types of Assessment

- Classroom assessments
  - For diagnosis
  - For grading
- State-wide end-of-course tests
  - Including AP like assessments
- Assessments linked to grade-level standards
- Nation-wide assessments
  - With national norms
  - Linked to "generic" standards
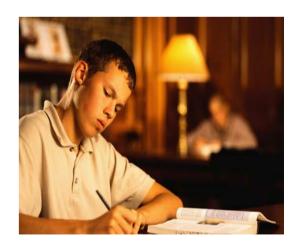- International assessments

- College entrance tests
  - General skills
  - Subject-specific tests
- College placement tests
- Work-readiness tests
- Vocational interest and guidance tests
- Personality tests

Overview of Key Technical Issues

➢ Technical Issues for All Assessments

   ▪ Validity – Does the assessment measure what we want it to?

      ✓ Clarity of Content Specifications

      ✓ Alignment Studies

   ▪ Reliability – How accurate are the resulting scores?

      ✓ Classification or Decision Accuracy

   ▪ Diagnostic Utility – Do results indicate strengths and weaknesses?

   ▪ Accessibility for English Learners and Students with Disabilities

➢ Particular Concerns for High School Assessments

   ▪ Participation and Effort

   ▪ Curricular/Instructional Validity

   ▪ Accounting for Dropouts

Participation and Effort

➢ The main barrier to state-level 12th grade assessments for NAEP is perceived problems with participation and effort.

- Even if school participation is made mandatory, student participation is not; recent rates have been dangerously low.

- Students may be unavailable for a variety of reasons, ranging from frequent absences to participation in co-enrollment programs.

- Results have absolutely no consequences for students and schools.
  - ✓ Confidentiality is a key concern for both students and schools.
  - ✓ Due to matrix sampling, scores for individual students and small schools would have low reliability.

- Myth and anecdote concerning a "senior slump" are hard to overcome with data.

Participation and Effort (Continued)

➢ Also a problem for some state assessments.

- Not perceived as a problem if results have consequences for the students, such as:
  - ✓ Classroom or end of course tests that affect student grades
  - ✓ Graduation or college entry tests
- Participation problems generally solved for census tests used for NCLB.
- Student effort on NCLB tests with no consequences for students is an open question.
  - ✓ Seniors know what "voluntary" means (unfortunately).

Possible Solutions to Participation and Effort Concerns

- ➢ Increasing Participation
  - ▪ Provide incentives to schools for high student participation.
    - ✓ This is how GED gets participation in norming studies.
    - ✓ NCLB provides negative incentives.
- ➢ Assuring Effort
  - ▪ Measure it! (e.g., fit indices, quitting, ref. Cohen)
  - ▪ Provide incentives for schools for high levels of student effort.
    - ✓ Different incentives work for different students.
  - ▪ Introduce (possibly positive) consequences for students
    - ✓ Seals and certificates, competitions among classes or schools.

Curricular/Instructional Validity

- ➤ What this really means is:
  - ■ Does the student's curriculum cover the knowledge and skills being assessed?
  - ■ Is instruction in this curriculum adequate?
- ➤ What it could also mean is:
  - ■ Does the assessment have a positive impact on curriculum and instruction?
- ➤ These are legal issues if the assessment has important consequences for students!

Curricular/Instructional Validity

➢ These issues are particularly salient in high school assessments because:

- Different students take different courses.
- There may also be alternative versions of a single course
  - ✓ Two-year and one-year Algebra I courses
- Teachers are generally expected to have credentials for the subject(s) they teach.
- English learners and many students with disabilities are further behind and their courses are less likely to cover the general education curriculum.
- Graduation tests trigger extreme legal scrutiny with respect to these issues.

Options for Dealing with Curriculum Diversity

➢ Option 1: Only test material all students should master.

- Advantages:
  - ✓ Allows all students to be included.
  - ✓ Focus on whether students are being left behind.
- Disadvantages:
  - ✓ Ignores high school contributions through elective courses.
  - ✓ May lead to focus on "pre-high school" content (or minimal competencies)
  - ✓ Still requires evidence that all students have received adequate instruction in this material.
  - ✓ Timing questions if instruction follows or significantly precedes the assessment.

Options for Dealing with Curriculum Diversity

➢ Option 2: Only test students completing specific courses.

- Advantages:
  - ✓ Enables tight correspondence between assessment and curriculum.
  - ✓ Allows inclusion of higher-level courses.
- Disadvantages:
  - ✓ Might promote "gaming" the system by restricting who can take critical courses.
  - ✓ Might fail to identify students being left behind because they do not take the course.
  - ✓ Difficulties in comparing proficiency across different courses.
  - ✓ Students might take course prior to high school.

➢ Hybrid Example: Virginia Graduation Requirement

- Students must pass a fixed number of required and elective end-of-course tests.

Options for Dealing with Curriculum Diversity

➢ **Option 3: Assess "Readiness": What all students need for success in college, job training, and entry into the military.**

- Advantages:
  - ✓ Content can be based on predictive rather than curricular validity – leave it to districts and schools to figure out how to develop the skills.
  - ✓ Likely to include skills (e.g., higher order thinking) beyond minimum competencies.
- Disadvantages:
  - ✓ Difficulty in getting agreement on readiness skills:
    - Limited to verbal and quantitative reasoning?
    - Include skills such as teamwork or listening skills or academic learning skills?
    - Predictive validity studies can be difficult and time-consuming.
  - ✓ Still leaves out much of what high school currently offer:
    - Specialized content covered by elective courses.
- Watch 12th Grade NAEP for progress on this option!

Options for Dealing with Curriculum Diversity

➢ Case Study: NAEP efforts to assess Foreign Language

- Pilot Test covered Spanish only
- Two levels:
  1. Students with at least one year of instruction in Spanish
  2. Students with more than two years of instruction.
- Argument about inclusion of native speakers of Spanish.
  - ✓ Assessing impact of instruction versus ability of American youth to speak Spanish?
- Experts insisted on assessing listening and speaking as well as reading and writing in Spanish.
- Outcome of a Pilot Test:
  - ✓ Logistical issues in identifying appropriate students to test.
  - ✓ Reliability and logistical concerns in assessing listening and speaking.
  - ✓ Assessment of foreign language deferred indefinitely!

Dropouts

➢ What can we say about dropouts?

- Not generally feasible to survey/assess students who have already dropped out.

- Can we say anything about what they learned prior to dropping out?

- Less of a problem for 10th grade census tests, than for 12th grade assessments.

➢ NCLB requires separate accountability for graduation rates.

➢ Better data systems are needed!

- Privacy protection may not be beneficial to students if it means no one realizes that they need help.

What States are Doing

➢ See handouts for some information on how states are implementing high school accountability and, in some cases, graduation exams.

Audited Transcript Model for High School Assessment

➢ A New/Old Model: Use transcripts to assess student achievement during high school.

- Audit/adjust grades through end-of-course assessments.
  - ✓ Census or sample assessments
  - ✓ Similar to tests used to certify performance in AP courses.
  - ✓ Universities often adjust for differences in grading practices based on external test scores or first-year college success.
  - ✓ Audit feedback might also improve consistency in grading practices.
- The National Center for Education Statistics regularly conducts high school transcript studies.

Audited Transcript Model

> Some advantages:

- Comprehensiveness: covers all high schools academic content.
- Uses available information: transcripts already maintained for other uses.
- Provides information on particular subjects (percent taking and performance) as well as overall (g.p.a.)
- Includes dropouts.
- Grade metric familiar to policy-makers and the public (arbitrary assessment scales are not).

Audited Transcript Model

➢ Disadvantage: Details not yet worked out.

➢ Research Needs:

- How to value different transcripts.
  - ✓ Importance of different types of courses for readiness.
  - ✓ Adjusting for course difficulty and grading practices.
- How to report results.
  - ✓ Overall achievement
  - ✓ Subject specific achievement
- Logistical difficulties
  - ✓ Obtaining transcripts
  - ✓ In a common electronic form.

Table Discussion

➢ Time Permitting:

- Pick one or two needs of high school assessments that are important to those at your table.

- Suggest the best type(s) of assessments to meet these needs.

- Suggest models for dealing with the technical issues discussed here:
  - ✓ Participation and effort
  - ✓ Curricular diversity
  - ✓ Accounting for dropouts

Some Useful References

➤ Achieve. (n.d.). State service benchmarking criteria. Archived at
http://www.achieve.org/achieve.nsf/StateServices_Criteria?OpenForm

➤ American Federation of Teachers (2001). *Making standards matter 2001: A fifty-state report on efforts to implement a standards-based system*. Retrieved July 26, 2005, from http://www.aft.org/pubs-reports/downloads/teachers/msm2001.pdf

➤ Center on Education Policy. (2004). *State high school exit exams: A maturing reform.* Retrieved July 26, 2005, from http://www.cep-dc.org/highschoolexit/ExitExamAug2004/ExitExam2004.pdf

➤ Council of Chief State School Officers. (n.d.). State accountability system: High school assessments. Archived at
http://accountability.ccsso.org/search.asp?stateID=ALL&subCategoryID=26&results=1&surveyYear=2003

➤ Council of Chief State School Officers. (2004a). *Statewide student assessment to meet NCLB requirements: English/Language Arts, 2003 – 2005*. Retrieved July 26, 2005, from http://www.ccsso.org/content/pdfs/EngLangArtsAssess2003-05.pdf

➤ Council of Chief State School Officers. (2004b). *Statewide student assessment to meet NCLB requirements: Mathematics, 2003 – 2005*. Retrieved July 26, 2005, from http://www.ccsso.org/content/pdfs/MathAssess2003-05.pdf

➤ Education Week. (2005a). How *Education Week* graded the states. *Quality Counts 2005 No Small Change: Targeting Money toward Student Performance, 24 (17),* .p. 104. Retrieved July 26, 2005, from http://www.edweek.org/ew/articles/2005/01/06/17sos-how.h24.html

➤ Education Week. (2005b). Summary of grades by state. *Quality Counts 2005 No Small Change: Targeting Money toward Student Performance, 24 (17),* .p. 82. Retrieved July 26, 2005, from http://www.edweek.org/cgi-bin/texis/search?querystring=target+money+toward+student+performance&source=all&date=all

➤ U.S. Department of Education. (n.d.). Decision letters on each state's final assessment system under NCLB. Archived at http://www.ed.gov/admins/lead/account/nclbfinalassess/index.html