

Linking Subnational and International Assessment Data: The U.S. Experience

Val Plisko

National Center for Education Statistics (NCES)

October 2010

This presentation is intended to promote the exchange of ideas among researchers and policy makers. The views expressed in it are part of ongoing research and analysis and do not necessarily reflect the position of the U.S. Department of Education.

Overview

- Assessment context in United States
 - State assessments
 - National Assessment of Educational Progress (NAEP)
 - State-level interest in international benchmarking
- Leveraging NAEP: NAEP-TIMSS 2011 Linking Study
 - Comparison of NAEP and TIMSS
 - Study design

State Assessments

- States required to test students in reading and mathematics
 - Annually in grades 3 through 8
 - At least once between grades 10 and 12
- States required to participate in national assessment (NAEP) in reading and mathematics with a sample of students in grades 4 and 8 every other year
 - Benchmarking purpose
- States, districts, and schools administer additional tests

State Assessments

- State assessments are uneven, in terms of...
 - Content
 - Skills
 - Difficulty and states' definitions of proficiency
- State-level NAEP results...
 - Enable comparisons of progress across states
 - Help states benchmark with other states and the nation

About NAEP

- National Assessment of Educational Progress (the “Nation’s Report Card”)
 - largest U.S. nationally representative and continuing assessment
 - Congressionally mandated
 - Data at national and state levels, as well as more than 15 large urban school districts
 - Results for subgroups (sex, race/ethnicity, student and school-level socioeconomic status)

About NAEP (continued)

- NCES responsible for administration of NAEP
- National Assessment Governing Board (NAGB) establishes frameworks and benchmarks
 - Independently sets achievement levels at which U.S. students should perform
 - NAGB members include subject-area experts, teachers, policymakers, business representatives, and members of the general public

Main NAEP Components

- Grades 4, 8, and 12
- National level and state level (since 1990) for mathematics, reading, science, and writing, usually grades 4 and 8
- Other subjects include (national only): the arts, civics, economics, geography, and U.S. history
- At national level, includes public and private schools
- At state level, public schools only

NAEP Scale Scores and Achievement Levels

- Achievement reported in two ways—scale scores and achievement levels
- Three achievement levels at each grade:
 - *Basic*--partial mastery of prerequisite knowledge and skills fundamental for proficient work at each grade
 - *Proficient*--solid academic performance and competency over challenging subject matter
 - *Advanced*--superior performance

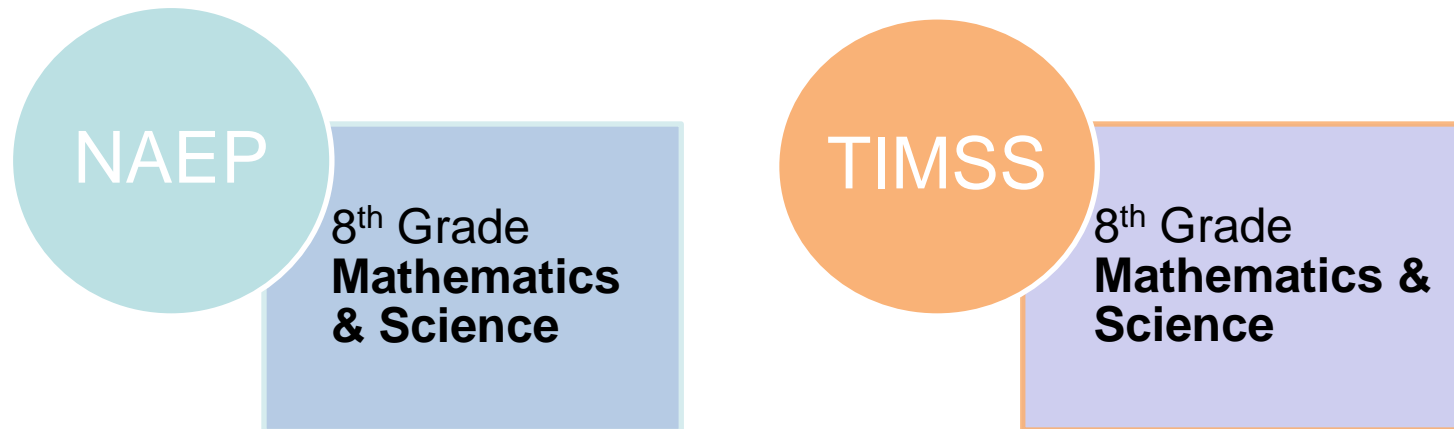
New Dimension – Benchmarking “Standards”

- State interest in benchmarking standards
- One method of benchmarking standards is to compare assessments:
 - How does “basic” or “proficient” or “advanced” in my state compare with “low” or “intermediate” or “advanced” in other countries?
 - What percentage of students in my state reach various international benchmarks? How does that compare with other countries?
- Federal interest in common standards that are higher, clearer, fewer, and internationally benchmarked
 - “Race to the Top” and ARRA (2009)

States and TIMSS

- U.S. states have drawn representative samples and obtained TIMSS scores
 - Most recently, Massachusetts and Minnesota (2007) and Indiana (2003)
- How can we support state interest in TIMSS without dramatically increasing burden to schools and students and cost to states?
- How can we take advantage of existing data collections?

NAEP-TIMSS 2011 Linking Study: Goals of the Study



- To project state and school district TIMSS mathematics and science scores from performance on NAEP
- To validate projected state TIMSS scores with real TIMSS scores in several states

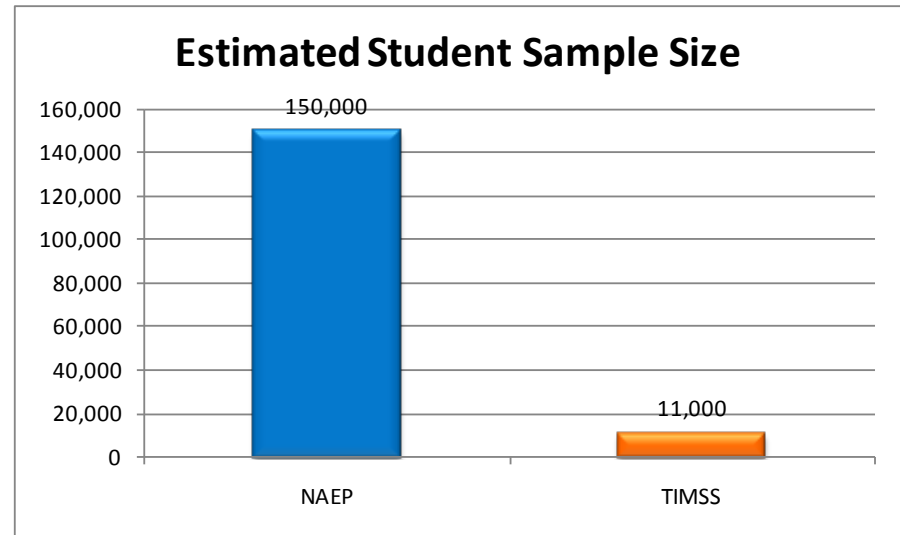
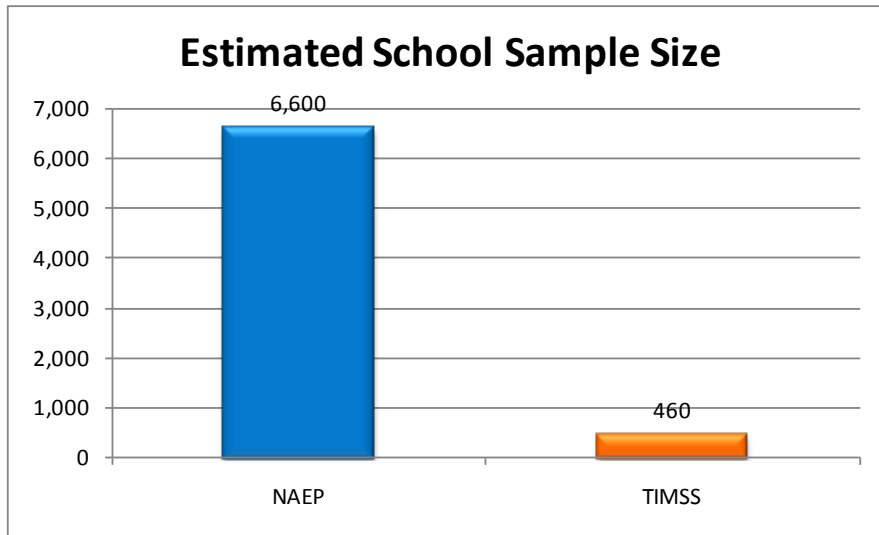
Comparing NAEP (2005/07) and TIMSS (2007)

- Some NAEP-TIMSS similarities
 - ✓ Both assess mathematics and science at grades 4 and 8
 - ✓ Test administration at similar (sometime overlapping) time periods
 - ✓ Both aim to measure school-based curricular attainment
 - ✓ Mathematics frameworks organized similarly
 - ✓ Over 95 percent of TIMSS mathematics items from both grades could be placed in the NAEP mathematics framework
 - ✓ For mathematics, 86 percent of grade 4 TIMSS items were consistent with topics in the grade 4 NAEP framework
 - ✓ Trend results for mathematics at grades 4 and 8 are consistent—for students overall and for boys and girls separately

Comparing NAEP (2005/07) and TIMSS (2007)

- Some NAEP-TIMSS differences
 - ✓ Science frameworks differ somewhat in organization, and cognitive dimension defined differently for both mathematics and science
 - ✓ Greater emphasis placed on *measurement* in NAEP and on *geometry* in TIMSS at grade 4, and a greater emphasis was placed on *geometry* in NAEP and on *number* in TIMSS at grade 8
 - ✓ NAEP placed a greater emphasis on *Earth science* at both grades; conversely, TIMSS placed a greater emphasis on *life science* at grade 4 and on *physical science* at both grades
 - ✓ For science, 79 percent of grade 4 TIMSS items were consistent with topics in the grade 4 NAEP framework; 56 percent of grade 8 TIMSS items were consistent with the grade 8 NAEP framework
 - ✓ Higher percentage of multiple choice in NAEP in math

Comparing NAEP and TIMSS: School and Student Projected 2011 Sample Sizes



NAEP-TIMSS 2011 Linking Study: Goal 1

- To project state TIMSS mathematics and science scores from performance on NAEP
 - ✓ Two sets of “braided” booklets containing both NAEP and TIMSS grade 8 mathematics and/or science items will be given to samples of eighth-grade students, one during regular NAEP window and other during regular TIMSS window
 - ✓ Data from these two sets of booklets will be used to link NAEP and TIMSS by estimating a function that predicts TIMSS scores based on state NAEP performance

NAEP-TIMSS 2011 Linking Study: Goal 2

- To validate projected state TIMSS scores with real TIMSS scores in several states
 - ✓ TIMSS state benchmarking will be administered in eight states.
 - ✓ Benchmarking consists of a normal TIMSS assessment of mathematics and science administered to state samples.
 - ✓ Data from TIMSS state benchmarking will be used to validate accuracy of projected state TIMSS score distribution estimated as a function of state mathematics and science performance on NAEP.

NAEP-TIMSS 2011 Linking Study: Study Approach

1. Administer braided booklets in both testing windows
 - NAEP in Winter (January to March, 2011)
 - TIMSS in Spring (April to June, 2011)
2. Select national samples independent of main NAEP and TIMSS samples
 - About 10,000 students in NAEP in Winter
 - About 11,000 students in TIMSS in Spring
3. Collect data using NAEP field staff in winter and TIMSS field staff in spring (same contractor for both)
4. Score items using same scoring staff , training, and QC procedures that NAEP and TIMSS use

NAEP-TIMSS 2011 Linking Study: NAEP Study Components

- NAEP assesses random samples of students within schools
- About half of the 10,000 students will be assessed in math and the other half in science
- Braided test booklets administered during the NAEP window will contain:
 - ✓ one 25-minute cognitive block from NAEP
 - ✓ one 25-minute cognitive block from TIMSS
 - ✓ two sets of NAEP student background questions taking total of 15 minutes

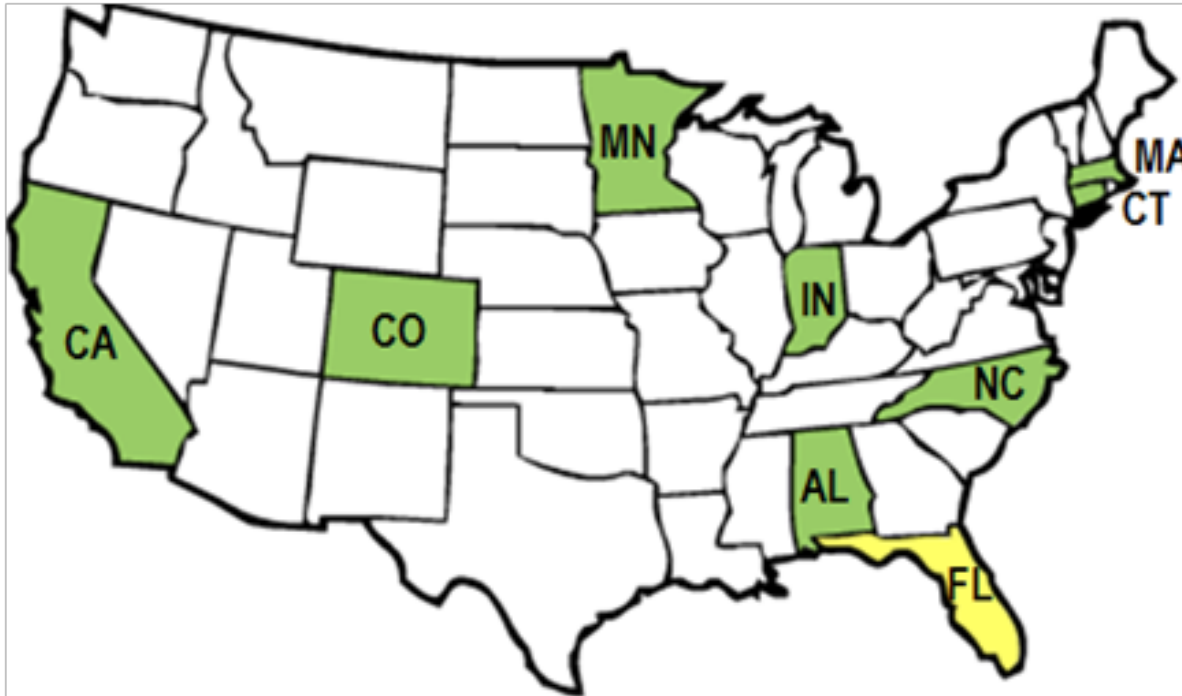
NAEP-TIMSS 2011 Linking Study: TIMSS Study Components

- TIMSS program assesses samples of intact classrooms
- This study sample will be drawn from same schools as main TIMSS sample, but different classrooms
- All 11,000 students will be assessed in math and science
- Braided test booklets administered during TIMSS window will contain three separately timed sections:
 - ✓ one NAEP and one TIMSS block timed at 47.5 minutes in same subject (math or science)
 - ✓ two TIMSS blocks timed at 45 minutes, one math and one science
 - ✓ one 30-minute section of student background questions

NAEP-TIMSS 2011 Linking Study: Selection of Validating States

1. States eligible for selection:
 - Large enough population
 - Agree to participate
 - 29 states eligible for selection
2. Selection criteria:
 - Prior NAEP performance (highest priority criterion)
 - ✓ Include states that represent both ends of NAEP achievement range to better test linking function
 - Prior participation in TIMSS (to compare trends in NAEP and TIMSS)
 - Regional representation (two states from each region: Midwest, Northeast, South, and West)

NAEP-TIMSS 2011 Linking Study: Selected States



Alabama, California, Colorado, Connecticut, Indiana, Massachusetts, Minnesota, and North Carolina

NAEP-TIMSS 2011 Linking Study: Benefits of the Study Approach

- Avoids interference with the conduct of regular operational assessments
- Provides a NAEP-TIMSS link through common items taken by the same student
- Measures the amount of growth during the period from winter to spring
- Provides states with international benchmarks in mathematics and science at grade 8, without incurring cost of conducting TIMSS

Conclusion

- NCES trying to help states fulfill international benchmarking goals:
 - Without dramatically increasing cost or burden to schools and students already doing lots of testing
 - Without doing harm, and possibly with some benefit, to U.S. participation in international assessments
- Stay tuned