GOALS AND CHOICES

A DISCUSSION OF:
OVERVIEW OF ILSA DESIGN ISSUES
IMPLICATIONS OF COMPUTER BASED TESTING

Ina V.S. Mullis
Executive Director, TIMSS & PIRLS International Study
Professor, Educational Research, Measurement, Evaluation
Lynch School of Education
Boston College

National Academy of Education
Workshop on International Large-Scale Assessment
June 17, 2016
One Size Does Not Fit All ILSAs

- Different Goals
- Different Clients
- Different Resources
- Different Choices
- Different Methods
TIMSS and PIRLS

- Designed by researchers for school improvement – not accountability
- Trend studies over 20 years
  - Evolving design (no more than 3 cycles)
  - Bridging
- Assess what countries expect students to learn
- Extensive array of background questionnaire data
TIMSS and PIRLS Content Frameworks

- Grades 4 and 8
- Based in countries’ official curricula
  - What students are expected to know and be able to do
- Agreed upon by participating countries
  - Stabilized by being trend studies
- Initiatives country based
Tailoring Items

Agree completely – assessment difficulty should match the countries’ capabilities

- TIMSS and PIRLS do not have a ceiling effect
- TIMSS Numeracy
- PIRLS Literacy
- LaNA
- Convincing countries to participate in the most appropriate assessments
Improving Measurement of Key Reporting Variables

Agree completely – questionnaire data susceptible to measurement error

- Working toward indices of 8-12 items to measure policy relevant constructs
  - e.g., student engagement, school safety
- Increase reliability
- Measure trends
Answering Causal Questions

Square pegs in round holes – ILSAs not optimal for causal analysis, should concentrate on strengths

- Context Questionnaire Framework
  - Literature about effective practices
  - Information the countries want collected
Longitudinal Data in TIMSS

Re-occurring idea

- TIMSS has offered to help countries
- No country has followed through
- Cost/benefit perspective
  - Huge burden for countries
  - Modest gains in analysis/interpretation
Computer-Based Testing

Promises, promises…but it is the future

- Better measurement
- More engaging, motivating
- Operational efficiencies
- Process data to inform instruction
Why Do Countries Want CBT?

- ePIRLS—online reading
  - Students reading online for school
  - Reading assessment should assess online reading
  - Conducted on voluntary basis
- eTIMSS—various reasons
  - Cheaper? (same as paper but delivered digitally)
  - Testing format more consistent with students’ lives?
  - Everybody else is doing it?
Extremely Labor Intensive to Realize Potential

It takes a village…

- Content and measurement specialists
- Graphic artists, front-end developers, back-end developers
- New assessment paradigms
  - Learning through the assessment
  - Process data
- Exciting and exhausting
Digital Divide

- Worldwide many students without computer/internet access
- A number of TIMSS and PIRLS countries do not have the necessary infrastructure
- eAssessments and paper assessments will grow further and further apart as we improve our eAssessments
- Mobile technologies – new challenges and opportunities
THANK YOU!

Ina V.S. Mullis
Executive Director, TIMSS & PIRLS International Study
Professor, Educational Research, Measurement, Evaluation
Lynch School of Education
Boston College

National Academy of Education
Workshop on International Large-Scale Assessment
June 17, 2016