Developing Assessment Frameworks for Measuring Trends in a Changing Context

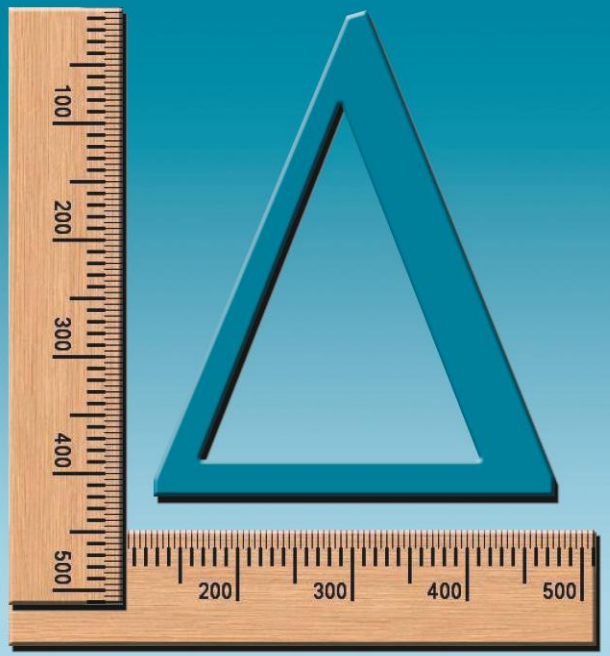
Ina V.S. Mullis and Michael O. Martin

Executive Directors
TIMSS & PIRLS International Study Center
Boston College

53rd IEA General Assembly
October 2012, Phuket

Updating TIMSS and PIRLS Frameworks for the Next Assessment

- Very different from developing a framework for a new assessment of a single subject in an individual country
- TIMSS and PIRLS are well established programs for measuring trends, with histories and constraints
- Goal to improve each assessment!!



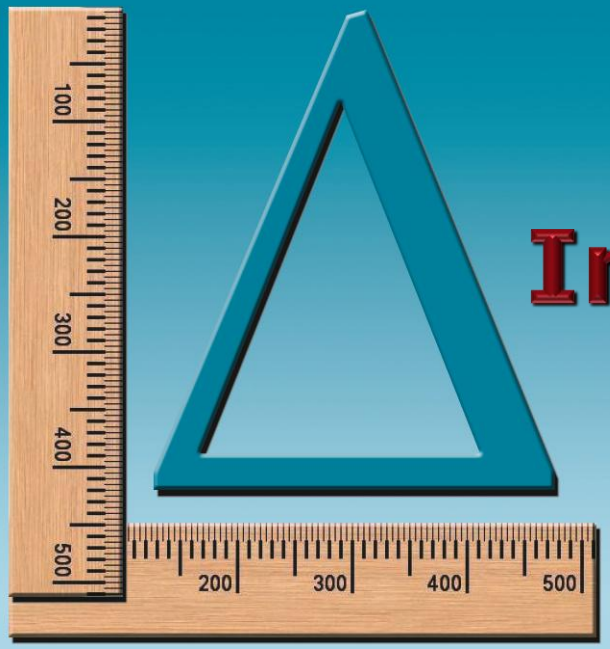
Important Principle

When measuring change,
do not change the measure.

Albert E. Beaton
John W. Tukey

Maintaining Relevance also Crucial Goal

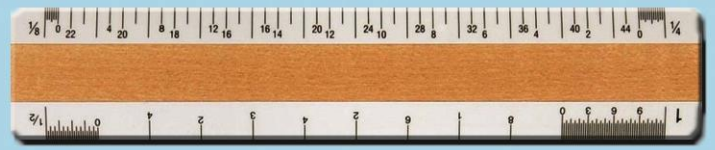
- Our world is constantly changing
- Countries rely on TIMSS and PIRLS for monitoring progress in educational achievement
- Must be relevant to current learning goals
- Must address current policy issues



Addendum to Important Principle

When measuring change in a changing world, it is important to change the measure.

Ina V.S. Mullis
Michael O. Martin



Seemingly Conflicting Goals

#1 Do not change

TIMSS and PIRLS are trend studies and cannot change dramatically

#2 Keep current

TIMSS and PIRLS must stay at the forefront of providing high quality measures

Need to Evolve with Each Assessment Cycle

TIMSS and PIRLS resolve tension between

- Maintaining continuity with the past procedures
- Maintaining current relevance in a changing context

Specific design to steadily replace items with each new assessment

Evolving Design

The evolving design used in TIMSS and PIRLS

- 60% items retained, 40% new items
- After three cycles for TIMSS and four cycles for PIRLS, all items are released and replaced with new
- For TIMSS 2011, all 1995 and 1999 items released
  - 20% from 2 cycles ago (e.g., 2003)
  - 40% from 1 cycle ago (e.g., 2007)
  - 40% new for 2011

Keep Present as Point of Reference

- Link backwards while moving forwards
- Keep substantial portions of assessment constant (e.g., 3 literary and 3 informational passages in PIRLS)
- Introduce new aspects carefully and gradually (e.g., 2 literary and 2 informational passages)
- Plan as trend assessment

Frameworks Need to Be Forward Looking

- Steadily improve each assessment
- More useful information
- Improved measurement
- More efficient data collection

Making International Comparisons: Another Important Goal and Constraint

- TIMSS and PIRLS comparative studies
  - Not a single country but MANY countries
- Assessment content the same across all countries
- Conducted in a uniform manner across all countries
- Necessary for assessments to be feasible and applicable for all participating countries

TIMSS and PIRLS Assessment Frameworks

- Framework for measuring achievement in the subject – mathematics, science, reading
- Framework for measuring contexts for teaching and learning in the subject
- Assessment block/booklet design
  - Assessment scope and student testing time
  - Item types

Updating the Assessments

- Collaborative process among countries
- National Research Coordinators of participating countries provide new ideas representing their countries' point of view
- Participating Countries
  - Expend the resources
  - Do an enormous amount of work
  - Entitled to influence the decisions

National Research Coordinators

- Designated by participating countries
- Responsible for implementing studies in their countries
- Involves responsibility for implementing various complex measurement and assessment tasks
- Very experienced and knowledgeable group

Key Participants in Updating Frameworks

- NRCs – ideas for improvement
- TIMSS & PIRLS International Study Center
  - Ensuring the quality of the study
  - Responsible for keeping studies within constraints
  - Leadership in meeting the data needs of as many countries as possible
- Chief Consultants
  - TIMSS Mathematics: Liv Sissel Gronmo, Norway
  - TIMSS Science: Lee Jones, United States
  - PIRLS Reading: Marian Sainsbury, England

Reviewers/Advisors

- International Expert Committees
  - TIMSS Science and Mathematics Item Review Committee (SMIRC)
    - Mathematics
    - Science
  - PIRLS Reading Development Group (RDG)
- Nationally or internationally recognized experts in their fields
- Promote new directions in the field
- Guarantee the accuracy of the subject matter content

Reviewers/Advisors

- Representative of participating countries
  - Geographic
  - Cultural
  - Economic
  - Language (reading)
  - Gender balance
  - Subject specialty (science)
  - Grade specialty (4, 8, 12)
- Approx. 8-10 experts per subject

Process for Updating Frameworks

- 1st NRC meeting discuss how to update assessments and frameworks
  - NRCs propose improvements
  - TIMSS & PIRLS International Study Center works with Chief Consultants and Expert Committees to propose updates
- TIMSS & PIRLS ISC
  - Conducts online survey
  - Drafts updated Framework
- Draft reviewed by Expert Committees
- 2nd NRC meeting review and agree Framework

Framework Development Ongoing Process

- TIMSS sixth cycle
- PIRLS fourth cycle

Many NRCs involved for multiple cycles

- Collaboratively make decisions about assessment content and procedures
- Begin making proposals well in advance of first meeting

Updating Framework Part of Larger Assessment Process

Need coherence between essential components of assessment framework

- Description of the assessment goals
- Blueprint for item development

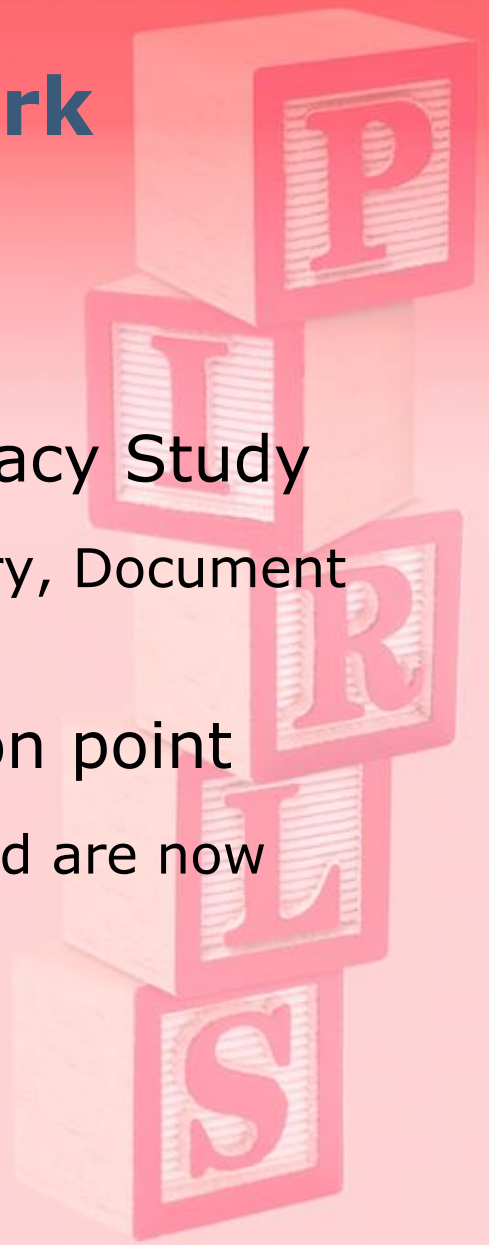
Need Framework that can be assessed

- Resources for implementation
- Feasibility of implementation across countries

History of PIRLS Framework Development

PIRLS 2001 Framework

- Building on IEA 1991 Reading Literacy Study
  - Was organized by Narrative, Expository, Document text types
- Fourth grade an important transition point
  - Students have learned how to read and are now reading to learn
  - Matched TIMSS



History of PIRLS Framework Development

- Identified Expert Committee
  - Reading Development Group, the RDG
- Addressed best thinking about skills and strategies needed for high degree of reading comprehension
  - Reading literacy as a constructive and interactive process
- Articulated home and school contexts for learning to read

History of PIRLS Framework Development –cont.

- Updated definition of reading comprehension to include
  - Construct meaning from a variety of texts
  - Read to learn, to participate in communities of readers, and for enjoyment

History of PIRLS Framework Development –cont.

- Reconceptualized the text type organization
  - Purposes for reading – literary and informational
  - Processes of comprehension
- Emphasized authentic assessment
  - Longer passages from authentic sources
  - Constructed-response questions requiring analysis supported by textual evidence
  - The *PIRLS Reader* in color and magazine format

History of PIRLS Framework Development –cont.

PIRLS 2006 Framework

- Primary NRC recommendation – enable reporting by comprehension processes
- Expanded assessment
  - Increased number of assessment passages and items
  - Combined four comprehension processes into two new process scales for reporting
  - Based on research studies in Germany and the TIMSS & PIRLS ISC at Boston College

History of PIRLS Framework Development –cont.

- Enhanced the description of the range of texts required to assess and report comprehension processes
- Elaborated description of informational reading to better describe the wide variety of texts
- Expanded Contextual Framework to explicitly address classroom instruction and experiences

History of PIRLS Framework Development –cont.

PIRLS 2011 Framework

- Primary NRC recommendation – update and expand the Contextual Framework
  - Improve validity and reliability by developing more robust, policy relevant scales
  - Relationship to achievement
  - Instructional effectiveness – student engagement

History of PIRLS Framework Development –cont.

- Primary NRC recommendation – address web-based reading
  - Expanded view of text types to encompass information and communication technologies (e.g., internet, email, and video)
  - Developed a prototype for assessing informational reading in a web-based as well as paper and pencil format

History of PIRLS Framework Development –cont.

- Extended the PIRLS assessment to be more inclusive of countries where many students find PIRLS too difficult, while preserving measurement quality
  - Difficulty of the assessment should match the ability of the students
- Option to assess PIRLS at higher grades (5^{th}, 6^{th})
- prePIRLS – a less difficult version of PIRLS
  - Makes it possible for a range of developing countries to assess reading at the end of the primary school cycle

History of PIRLS Framework Development –cont.

Updating PIRLS 2016 Framework

- Invitations and agenda for 1st NRC meeting distributed to NRCs – Hamburg, Feb 2013
  - Update descriptions/definitions of purposes and processes (especially informational purpose)
  - Update Contextual Framework
  - Refine criteria for submitting/reviewing assessment passages, including text complexity
  - Assess Web-based reading
  - Other initiatives

History of TIMSS Framework Development

TIMSS 1995 Frameworks (Also TIMSS 1999)

- *Curriculum Frameworks for Mathematics and Science*, Robitaille, et al., 1993
- Two aspects - Content and Performance Expectations
- A single content framework spanning the entire range of mathematics, and another science
  - e.g., *Numbers* – "Whole Number Operations" through "Complex Numbers and their Properties"
- Areas and topics listed but not elaborated

History of TIMSS Framework Development –cont.

TIMSS 2003 Frameworks

- Major update of content domains and topics
  - Additional support from U.S. National Science Foundation
- Supporting explanatory text describing and providing rationale for topics in each domain
- Assessment objectives specific to fourth and eighth grades
- New assessment design for measuring trends

History of TIMSS Framework Development –cont.

- Reconceptualized performance expectations as cognitive domains
  - Supporting explanatory text
- Expanded assessment size
  - E.g., from 160 to 215 items for 8th grade mathematics
- Calculators permitted for newly-developed assessment items at eighth grade (although not necessary to answer the questions)
  - Research design to monitor effect of calculator use

History of TIMSS Framework Development –cont.

TIMSS 2007 Frameworks

- Separate frameworks for fourth and eighth grades to facilitate item development
  - Grade specific content domains and topics
  - E.g., physical science at 4th grade; chemistry and physics at 8th grade
- NRC recommendation: Align cognitive domains across grades and subjects
  - Knowing, applying, and reasoning
  - Different levels of emphasis

History of TIMSS Framework Development –cont.

TIMSS 2011 Framework

- TIMSS and PIRLS together
- Update and expand the Contextual Framework
  - Early numeracy activities (PIRLS home questionnaire)
- Updated content and cognitive domains
  - Weightings, topics, and objectives
- Considered computer-based administration for TIMSS

History of TIMSS Framework Development –cont.

Updating TIMSS 2016 Framework

- Invitations and agenda for 1st NRC meeting distributed to NRCs – Hamburg, Feb 2013
  - Update content and cognitive domains (weightings, topics, and objectives)
  - Update Contextual Framework
  - Discuss learning progressions
  - Computerized assessment
  - preTIMSS
  - Other initiatives

In Summary…

Updating TIMSS and PIRLS assessment frameworks:

- Must evolve gradually to protect trends
- Must be forward looking to be relevant
- Must consider resource realities
- Must lead to an improved assessment
- Must be of the highest quality

Developing Assessment Frameworks for Measuring Trends in a Changing Context

Ina V.S. Mullis and Michael O. Martin

Executive Directors
TIMSS & PIRLS International Study Center
Boston College

53rd IEA General Assembly
October 2012, Phuket