Using TIMSS Data for Appraising the Impact of U.S. Math Reforms on Student Achievement: Additional Issues and Constraints

International Association for Evaluation of Educational Achievement
2nd IEA International Research Conference
November 9-11, 2006
Washington, DC USA

Daniel B. Berch

Overview of Presentation

- Limitations of this study
- Another TIMSS analysis using the RSM approach
- Issues and constraints based on ambiguities associated with reform-oriented practices:
  - Definitional complexities of “number sense”
  - Testing the effectiveness of cooperative learning

Acknowledged Limitations, Qualifications Constraints, Provisos, and Disclaimers

The field has not reached consensus on how to define "reform-oriented" or "traditional" instruction, so while there is extensive overlap across studies in how these constructs are conceptualized and measured, they do not always mean exactly the same thing.

The research literature suggests that only a small minority of U.S. teachers embody the principles of reform-oriented curriculum and instruction in their lessons.

TIMSS data do not allow researchers to follow individual students over time, which severely limits the range of analyses that can be conducted and the strength of the conclusions that can be drawn from these analyses.

Limitations, etc.

TIMSS surveys include a somewhat sparse set of items that directly address reform practices, changes in the wording of items from one wave to the next, and the strong possibility of incomparability of survey responses across countries.

The student survey items addressing reform-oriented practices did not cluster as clearly or consistently as the teacher survey items. These differences suggest that students in different countries might be interpreting the items differently.

In interpreting any relationships observed there it should be remembered that the precise *meaning* of attending a classroom where students report more or less of a certain activity may not be the same across countries.

Limitations, etc., continued

It should be noted that the alpha reliability estimates for *traditional perceptions* were generally low across countries, indicating that these composites may be less consistently measured (and thus less meaningful) than would be desirable.

The scales used to report coverage of topics also suffered some modifications. Thus, in the 1999 and 2003 surveys we cannot tell whether a topic is not covered at all during the year, and in 2003 we cannot separate between *not taught* and *just introduced*; etc.

The nature of the data collected by TIMSS (i.e., cross-sectional, non-experimental design with no prior-year scores for students) does not support causal interpretations.

Limitations, etc., continued

It is impossible to fully understand cross-country differences in achievement without considering the many contextual and cultural factors that can influence the academic experiences of students in classrooms.

Thus a truly accurate measure of a student’s exposure to instructional practices is very difficult to obtain with traditional survey methods and with the resource constraints typical of large-scale cross-sectional data collection efforts.

However closely we think teachers’ or students’ survey responses resemble their actual activities in the classroom, most surveys fail to provide any information on the *quality* of the practices in which teachers (and students) engage.

Limitations, etc., continued

Even when supplemented with information about curriculum or topics taught, most surveys fail to capture information about the extent to which teachers are implementing the core principles of reform, such as a focus on students' thinking.

Evaluations of specific curricula often fail to take into account the ways in which reform-based curricula can become distorted in practice, so that the instruction students receive fails to match the goals of the curriculum developers.

These (the TIMSS) assessments were never validated for the purposes of detecting effects of instructional practice, and there is evidence of cross-country differences in their psychometric properties.

Rule-Space Analysis of TIMSS-R, 1999 Content and Process Skills for Eighth Grade Students

Tatsuoka, Corter, & Tatsuoka (2004)

Rule-space method: Converts performance on specific test items for each student into an output vector of attribute mastery probabilities. The authors measured 23 specific content knowledge and cognitive processing subskill *attributes* (defined and validated in advance) underlying students' item scores.

Standardized Means of Conceptually-Based, Composite Achievement Variables

(After Tatsuoka et al., 2004)

Country	F1	F2	F3
Singapore	.45	.37	.70
Japan	.31	.51	.49
Netherlands	.32	.23	.13
USA	.11	-.33	-.07

F1 = Application, verbal, problem search, and *spatial skills*
F2 = Logical and abstract reasoning skills
F3 = Algebraic and data & complex process management skills
Note: Range of mean scores = .70 to -1.00 (i.e., where these scores represent "attribute mastery probabilities.")

TIMSS 2003 Content and Cognitive Domains

Exhibit 2: Target Percentages of TIMSS 2003 Mathematics Assessment Devoted to Content and Cognitive Domains by Grade Level

	Fourth Grade	Eighth Grade
Mathematics Content Domains		
Number	40%	30%
Algebra*	15%	25%
Measurement	20%	15%
Geometry	15%	15%
Data	10%	15%
Mathematics Cognitive Domains		
Knowing Facts and Procedures	20%	15%
Using Concepts	20%	20%
Solving Routine Problems	40%	40%
Reasoning	20%	25%

From Mullis et al. (2003)

Characteristics of TIMSS Number Content Domain

“The number content domain includes understanding of counting and numbers, ways of representing numbers, relationships among numbers, and number systems. At the fourth and eighth grades, students should have developed number sense and computational fluency, understand the meanings of operations and how they relate to one another, and be able to use numbers and operations to solve problems. The number content domain consists of understanding and skills related to:

- **Whole numbers**
- **Fractions and decimals**
- **Integers**
- **Ratio, proportion, and percent”**

TIMSS 2003 - Whole Number Understanding and Skills

Number: Whole Numbers

Grade 4

- Represent whole numbers using words, diagrams, or symbols, including recognizing and writing numbers in expanded form.
- Demonstrate knowledge of place value.
- Compare and order whole numbers.
- Identify sets of numbers according to common properties such as odd and even, multiples, or factors.
- Compute with whole numbers.
- Estimate computations by approximating the numbers involved.
- Solve routine and non-routine problems, including real-life problems.

Grade 8

- Demonstrate knowledge of place value and of the four operations.
- Find and use factors or multiples of numbers, and identify prime numbers.
- Express in general terms and use the principles of commutativity, associativity, and distributivity.
- Evaluate powers of numbers, and square roots of perfect squares to 144.
- Solve problems by computing, estimating, or approximating.

Precisely What Does “Number Sense” Mean?

Originally defined just over 50 years ago:

“A faculty permitting the recognition that something has changed in a small collection when, without direct knowledge, an object has been removed or added to the collection” (Dantzig, 1954).

Berch (2005) compiled a list of 30 alleged components of number sense:

Table 1

Alleged Components of Number Sense

1. A faculty permitting the recognition that something has changed in a small collection when, without direct knowledge, an object has been removed or added to the collection
2. Elementary abilities or intuitions about numbers and arithmetic
3. Ability to approximate/estimate
4. Ability to make numerical magnitude comparisons
5. Ability to decompose numbers naturally
6. Ability to develop useful strategies to solve complex problems
7. Ability to use the relationships among arithmetic operations to understand the base-ten number system
8. Ability to use numbers and quantitative methods to communicate, process, and interpret information
9. Awareness of various levels of accuracy and sensitivity for the reasonableness of calculations
10. A desire to make sense of numerical situations by looking for links between new information and previously acquired knowledge
11. Possessing knowledge of the effects of operations on numbers
12. Possessing fluency and flexibility with numbers
13. Can understand number meanings
14. Can understand multiple relationships among numbers
15. Can recognize benchmark numbers and number patterns
16. Can recognize gross numerical errors
17. Can understand and use equivalent forms and representations of numbers as well as equivalent expressions
18. Can understand numbers as referents to measure things in the real world
19. Can move seamlessly between the real world of quantities and the mathematical world of numbers and numerical expressions
20. Can invent procedures for conducting numerical operations
21. Can represent the same number in multiple ways depending on the context and purpose of the representation
22. Can think or talk in a sensible way about the general properties of a numerical problem or expression -- without doing any precise computation
23. Engenders an expectation that numbers are useful and that mathematics has a certain regularity
24. A non-algorithmic feel for numbers
25. A well-organized conceptual network that enables a person to relate number and operation propitious
26. A conceptual structure that relies on many links among mathematical relationships, mathematical principles, and mathematical procedures
27. A mental number line on which analog representations of numerical quantities can be manipulated
28. A nonverbal, evolutionarily ancient, innate capacity to process approximate numerosities
29. A skill or kind of knowledge about numbers rather than an intrinsic process
30. A process that develops and matures with experience and knowledge

Number Sense
Lower- and Higher-Order Features

Berch (2005)

Lower-order **perspective characterizes number sense as a** **<u>biologically-based</u>** **"perceptual" sense of quantity.**

Higher-order **view depicts it as an** **<u>acquired</u>** **"conceptual sense-making" of mathematics.**

<u>Lower-order features</u>: elementary intuitions about quantity, including the rapid and accurate perception of small numerosities and the ability to compare numerical magnitudes, to count, and to comprehend simple arithmetic operations.

<u>Higher-order features</u>: a deep understanding of mathematical principles and relationships, a high degree of fluency and flexibility with operations and procedures, a recognition of and appreciation for the consistency and regularity of mathematics, and a mature facility in working with numerical expressions.

Reform-Oriented Practices: Cooperative Learning

Reform-oriented mathematics classrooms focus on students being actively engaged in mathematical discourse in cooperative settings (NCTM, 2000).

“These reforms and many of the new curricula emphasized the importance of problem solving and inquiry, and were designed to promote increased use of instructional activities that reformers believed would promote students’ thinking skills. These activities included cooperative groups . . .” (Hamilton & Martinez)

Factors That Can Influence the Effectiveness of Cooperative Learning

Anderson, Reder, and Simon (2000) suggest that the effectiveness of the cooperative learning approach depends upon the degree to which it is structured or scripted. A 1994 National Research Council report on cooperative learning concluded that "a number of detrimental effects arising from cooperative learning have been identified--the "free rider," the "sucker," the "status differential," and "ganging up" effects" (p. 11). Meegan and Berg (2002) point out that the verbalizations of group members can interfere with the thought processes of other group members or impede them from elaborating upon their own ideas. In addition, the attitudes of the group members regarding collaboration can influence the effectiveness of their collaboration as well as their willingness to engage with others in a problem-solving situation.

Cooperative Versus Individual Learning: Rigorous Comparisons

Laughlin, Hatch, Silver and Boh (2006) compared group vs. individual performance using a letters-to-numbers coding problem, which combines facets of mathematical and logical reasoning as well as hypothesis testing. Rather than simply comparing a group's outcome with that of the average individual, these investigators used a more stringent test in which they examined multiple groups (*n*) of a given size (*m*) and an equivalent number of individuals (*nm*) (e.g., 40 groups of size 4 compared with 160 individuals), with subjects randomly assigned to these conditions.

This approach permitted a comparison of the mean of the consensus responses for each replication of a given-sized group with the mean of the best individuals' responses, second best, third best, etc., drawn from each of the replications (e.g., the mean consensus response of the 40 groups of size m = 4 compared with the mean of the 40 best individuals' responses, 40 second best, etc.).

Results of the Laughlin et al. (2006) Study and the Issue of Group-to-Individual Transfer

Laughlin et al. found that while groups of three, four, and five undergraduates did not differ from each other, they all performed better than an equivalent number of the best individuals (and all were better than groups of two).

But does group problem solving transfer effectively to individual students? That is, does it subsequently enhance individual problem solving skills? To date, there is little rigorous evidence pertaining to this issue. *Clearly, such transfer is crucial if individually-administered tests such as the TIMSS are going to be used for assessing the impact of reform-oriented instructional practices.*

Waschescio (1998) pointed out that while the constructivist perspective places great emphasis on social interaction, it rejects the idea that the shared meanings constructed from classroom interactions are internalized by the individual members of the group.

The U.S. National Mathematics Advisory Panel

"The National Mathematics Advisory Panel, created in April 2006 by President Bush to provide advice on the effectiveness of various approaches to teaching mathematics, includes individuals who have been associated with both sides of the debate. *Early indications of the panel's work suggest a greater degree of consensus than has been obtained in the past*" [emphasis added] (Hamilton & Martinez, 2006).

Unfortunately, some of the media coverage has raised questions and caused concern among our members. Despite several conversations with a reporter from the Wall Street Journal explaining what the Curriculum Focal Points are and are not, a September 12 Wall Street Journal article did not represent the substance or intent of the focal points. The focal points are not about the basics; they are about important foundational topics. The Council has always supported learning the basics. Students should learn and be able to recall basic facts and become computationally fluent, but such knowledge and skills should be acquired with understanding. Unfortunately, some of the other news media have followed the Wall Street Journal's lead and have similarly misrepresented the Curriculum Focal Points. The Council's goal is to support teachers in guiding students to learn mathematics with understanding.

What is Direct Instruction?

DI is comprised of explicit and systematic instructional formats based on scripted lesson plans, flexible skill grouping of students, brisk pacing of instruction, sequencing of skills, teaching to mastery, recurrent assessment, error correction, and the use of positive reinforcement. According to Huitt, Monetti, & Hummel (in press), devotees of DI programs take great pains to differentiate these from *teacher-made* lessons based on more generic models of direct instruction (di), mainly because the former have undergone rigorous standardization and field testing.

“Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching.”

Kirschner, Sweller, and Clark (2006)