DEVELOPING RIGOROUS AND MEANINGFUL LEARNING MEASURES FOR MIDDLE SCHOOL MATHEMATICS

Yvonne Kao & Jim Pellegrino

WestEd
University of Illinois at Chicago

BACKGROUND AND OBJECTIVES

The National Research and Development Center for Cognition and Mathematics Instruction (Math Center) is a large-scale effort to bridge research and practice by applying cognitive principles to redesign an existing middle school mathematics curriculum, Connected Mathematics Project 2 (CMP2), and testing the efficacy of these revised materials. The Math Center team selected cognitive-based principles shown to improve student learning: 1) integrating visual with verbal information to promote the integration of concepts, 2) structuring practice by interleaving worked examples and self-explanation prompts with new problems to solve, 3) carefully spacing the learning of critical content and skills over time, and 4) providing focused feedback on quizzes and homework.

The Math Center is conducting a full-year, cluster-randomized trial to test the efficacy of the 7th grade revised CMP2 materials. The efficacy study includes beginning-of-year and end-of-year standardized assessments of mathematics achievement as well as project-developed assessments for the end of each of the eight 7th grade CMP2 units. In this presentation we will discuss the process for developing these assessments.

Objectives

For each of the 7th grade CMP2 units, create a valid and reliable end-of-unit test that:

1. Is aligned to the content and practices emphasized in the unit
2. Can be compared to standardized measures of mathematics achievement
3. Is sensitive to the treatment manipulation

METHODOLOGY

Mapping the Content of the Curriculum

One of the central problems for developing a new assessment is identifying the concepts and skills that should be targeted. The teacher resources for CMP2 contain detailed information about the skills introduced, mastered, and reviewed in each unit. In addition, members of the UIC Math Center team carefully mapped the homework problems to determine the content and skills covered. CMP2 identifies a selection of homework problems for each unit as “core” problems. Worked examples were added to many of these core problems during the curriculum revision process. The core problems with worked examples thus represent the intersection between the most important content of the unit and the treatment manipulation and thus became the focus of the unit assessments.

Initial Item Selection and Calibration

Researchers sourced multiple-choice items for the item pool from nationally or internationally-validated, reliable tests (e.g., NAEP, TIMSS), from CMP2 materials, and open-ended items from the Balanced Assessment of Mathematics (BAM). All items in the item pool were mapped for content and skills.

Multiple-choice items for each unit’s field test were selected from this pool and arranged across four blocks: A, B, C, and D. Blocks were constructed to be of equivalent length (~9 items) and difficulty and so that each block contained content that was representative of the entire curriculum unit. Blocks were then paired in an overlapping block design to produce six test booklets, with each block appearing in three
different booklets so each item would be tested by 50% of the total calibration sample. In addition, each block appears both at the beginning of the test and at the end of the test in order to control for potential order effects.

For each unit, 2-3 BAM items were selected for the field test based on length and alignment to unit content.

Test booklets and BAM items were distributed for the field test such that each classroom took four of the test booklets and all of the BAM items. Test booklets were distributed so that approximately half of each class would take each block and each student received a randomly-selected booklet and a randomly-selected BAM item.

**Final Item Selection and Test Creation**
Selection of the final items for the assessments was informed by item response theory and constrained by the goals of the test design:

- Make one test form that can be completed in 40 minutes
- Include 1 item from the BAM (10-15 minutes)
- Include ~16 multiple-choice items (25-30 minutes)
- About half the multiple-choice items on each test form should be from standardized tests and the other half should be from CMP materials in order to ensure validity to CMP and comparability to broader research

Items were selected if they performed well statistically (e.g., good model fit, good point-biserial correlations, appropriate levels of difficulty) and aligned well to the content of the core problems with worked examples.

**ACKNOWLEDGEMENTS**

Many thanks to the following people for their contributions to this work: Deena Soffer Goldstein, Shandy Hauk, Aleata Hubbard, Anita Moorjani, and Cheryl Schwab.

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C100024 to WestEd. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.