Assessment Design

Assessment Design

Scope of the Assessment

To measure students' achievement in mathematics and science at the fourth and eighth grades and gather information about the contexts for achievement, the TIMSS 2003 assessment includes written tests of mathematics and science and a series of questionnaires focusing on contexts for student learning in those subjects. This chapter describes the design of the assessment and specifications for operationalizing the components of the study. A characteristic of TIMSS is that it includes both mathematics and science, with each student completing parts of the assessment in both subjects.

The TIMSS frameworks have broad coverage goals, and consequently the TIMSS Expert Panel found that a valid assessment of the mathematics and science described in the frameworks would require a substantial pool of assessment items and extensive testing time – at least seven hours at eighth grade (mathematics and science together) and more than five and one half hours at fourth grade. While the assessment material that can be presented in that time should provide good coverage of the mathematics and science students encounter at school and in their everyday lives, it is not reasonable to expect each student to answer the entire pool of test items.

Dividing up the Item Pool

Since the testing time required by the entire assessment item pool greatly exceeds the time available for testing individual students, TIMSS divides the assessment material among students. The TIMSS approach, based on matrix sampling techniques, involves dividing the item pool among a set of student booklets, with each student completing jpassessr at school

S14naticET .00 T j on1rwtat864 M 0 d GS3 gs407.635 329.130fm407.635 83r334 l478fl5 291.358fm478fl5 9

blocks will contain 15 minutes of assessment items and fourth-grade blocks 12 minutes; otherwise the general design is identical at both grade levels. The blocks containing mathematics items will be labeled M1 through M14 and the science items S1 through S14.

Since TIMSS in 2003 and in later cycles plans to provide an up-to-date assessment of student achievement in mathematics and science while also measuring trends in achievement since 1995 and 1999,[1] the TIMSS design for 2003 includes items from earlier assessments to measure trends as well as innovative new problem-solving and inquiry items and replacement items for those released into the public domain. Of the 14 item blocks in each subject, six (blocks 1 through 6) contain secure items from earlier TIMSS assessments to measure trends,[2] and eight (blocks 7 through 14) contain new replacement items.

Although calculators were not permitted in 1995 or 1999, calculators may be used in the eighth-grade assessment at the discretion of each participating country.[3] Calculators will not be permitted in the fourth-grade assessment.

Block Design for Student Booklets

In choosing how to distribute assessment blocks across student booklets, the major goal was to maximize coverage of the framework while ensuring that every student responded to sufficient items to provide reliable measurement of trends in both mathematics and science. A further goal was to ensure that trends in the mathematics and science content areas could be measured reliably. To enable linking among booklets, at least some blocks had to be paired with others. Since the number of booklets can become very large if each block is to be paired with all other blocks, it was necessary to choose judiciously among possible block combinations to keep the number of student booklets to a minimum.

The decision to allow calculator use at the eighth grade for the first time in 2003 also had an impact on the booklet design. Since calculators were not allowed in 1995 or 1999 but will be permitted in the eighth-grade assessment in 2003, it was necessary in order to safeguard the measurement of trend to arrange the item blocks in the booklets so that calculators could be used for the new assessment items, but not for the trend items. Accordingly, the trend blocks were placed in the first part of each booklet, to be completed without calculators before the break. However, two mathematics trend blocks (M5 and M6) and two science trend blocks (S5 and S6) also were placed in the second part of one booklet each.

In the TIMSS 2003 design, the 28 assessment blocks will be distributed across 12 student booklets (see Exhibit 7). The same booklet design will be used at both fourth and eighth grade, although the eighth-grade blocks will contain 15 minutes of assessment items and the fourth grade blocks 12 minutes. Each student booklet will consist of six blocks of items. Half the booklets will contain four mathematics

1 TIMSS ill mea re rend a he eigh h grade from 1995 and 1999, b a he fo r h grade from 1995 onl , ince he TIMSS 1999 a e men a cond c ed a he eigh h grade b no he fo r h grade.

2 The i rend block for he TIMSS 2003 eigh h-grade a e men ill con ain ec re i em from bo h he 1995 and 1999 TIMSS a e men , i h all rend i em from 1995 being placed in block 1 hro gh 3. Beca e here are no fo r h-grade i em from 1999, rend block 4 hro gh 6 for he fo r h-grade a e men ill con i of ne i em. In order o crea e balanced block , ome ne i em are al o incl ded in block 1 hro gh 6 a he eigh h grade.

3 To a oid in rod cing bia in o he mea remen of rend , calc la or ill be ed onl i h i em ne in 2003, and no i h i em mea ring rend .

In addition, scoring guides are designed to enable, for each item, identification of the various successful, partially successful, and unsuccessful approaches. Diagnosis of common learning difficulties in mathematics and science as evidenced by misconceptions and errors is an important aim of the study.

Since constructed-response questions constitute an important part of the assessment and are an integral part of the measurement of trends, it is very important for scoring guides to be implemented consistently in all countries and in each data collection year. To ensure consistent application of the scoring guides for trend items in the 2003 assessment, IEA has archived samples of student responses from each country; these will be used to train scorers in 2003 and to monitor consistent application.

Sc ▪e P . . In developing the assessment, the aim is to create blocks of items that each provide, on average, about 15 score points at eighth grade and about 12 score points at fourth grade. For example, at eighth grade blocks 1 through 14 in each subject could be made up of approximately 8 multiple-choice items (1 point each), 2 or 3 short constructed-response items (1 or 2 points each), and 1 extended constructed-response item (3 points). The exact number of score points and the exact distribution of question types per block will vary somewhat. Since the blocks for the fourth-grade assessment will be designed to yield 12 rather than 15 score points, there will be fewer items but the relative proportions of different item types will be approximately the same.

Scales for Reporting Student Achievement

TIMSS will report trends in student achievement in both the general areas of mathematics and science and in the major subject matter content areas. As each student will respond to only part of the assessment, these parts must be combined for an overall picture of the assessment results for each country. Using item response theory (IRT) methods,[4] individual student responses to the items related to mathematics and science will be placed on common scales that link to TIMSS results from 1995 and 1999. At the eighth grade, there will be an overall mathematics scale that will allow countries that participated in TIMSS in 1995 or 1999 to track their progress in mathematics achievement since then, and a similar scale in science overall that will provide the same information for science. At the fourth grade, the overall mathematics and science scales will link to 1995 only, since the TIMSS 1999 assessment did not include fourth grade. All students will have overall mathematics and science scores.

All student responses will contribute to the measurement of achievement in each of the mathematics and science content areas. In addition, those students assigned booklets with four blocks of mathematics items (half of the student sample) will provide the data to report on trends in mathematics content areas, while those assigned booklets with four blocks of science items (the other half) will provide data on trends in science content areas.

4 For a de▪crip ion of he TIMSS ▪caling echniq e▪ a▪ applied o he 1999 da a, ▪ee Yamamo o, K. and K lick, E. (2000), Scaling Me hod▪ and Proced re▪ for he TIMSS Ma hema ic▪ and Science Scale▪ in M.O. Mar in, K.D. Gregor , and S.E. S emler (ed▪.), *1 R* , Che▪ n Hill, MA: Bo▪ on College.

In mathematics at the eighth grade there will be five content reporting categories in 2003:

- Number
- Algebra
- Measurement
- Geometry
- Data

At fourth grade there also will be five content reporting categories in mathematics:

- Number
- Patterns, Equations, and Relationships
- Measurement
- Geometry
- Data

Eighth-grade science will have five content reporting categories:

- Life Science
- Chemistry
- Physics
- Earth Science
- Environmental Science

At fourth grade, science will have just three content reporting categories:

- Life Science
- Physical Science
- Earth Science

Results will be reported separately for each content area and grade level.

In addition to the IRT scales that will be used to summarize achievement in mathematics and science content areas and in these subjects overall, TIMSS will report on performance in each of the cognitive domains in terms of the average percentage of students answering items correctly in each domain. This approach may also be used to report student performance in scientific inquiry.

Releasing Assessment Material to the Public

The data collection in 2003 will be the third in the TIMSS series of regular four-year studies, and will provide data on trends in mathematics and science achievement since 1995 and 1999. TIMSS will be administered again in 2007, 2011, and so on into the future. The design provides for releasing many of the items into the public domain as the international reports are published, while safeguarding the trend data by keeping secure a substantial proportion of the items. As items are released, new items will be developed to take their place.

According to the TIMSS design, half of the 14 assessment blocks in each subject will be released when the assessment results for 2003 are published, and half will be kept secure for use in later assessments. The released blocks will include the three blocks containing trend items from 1995, one block of trend items from 1999, and three blocks of items used for the first time in 2003.[5]

and their school principals. The questions are designed to measure key elements of the curriculum as it is intended, as it is implemented, and as it is learned.

Curriculum Questionnaires. The curriculum questionnaires, one for mathematics and one for science, are designed to collect basic information about the organization of the mathematics and science curriculum in each country, and about the content in these subjects intended to be covered up to and including the fourth grade and between fourth and eighth grades. The National Research Coordinator in each country will be responsible for completing the questionnaires, drawing upon the knowledge and expertise of curriculum specialists and educators as necessary.

Student Questionnaire. This questionnaire will be completed by each student who takes the TIMSS assessment. It asks about aspects of students' home and school lives, including classroom experiences, self-perception and attitudes about mathematics and science, homework and out-of-school activities, computer use, home educational supports, and basic demographic information. The questionnaire requires 15-30 minutes to complete.

Teacher Questionnaires. In each school participating at the eighth grade, a single eighth-grade mathematics class will be sampled to take part in the TIMSS testing. The mathematics teacher of that class will be asked to complete a mathematics teacher questionnaire, providing information on the teacher's background, beliefs, attitudes, educational preparation, and teaching load, as well as details of the pedagogic approach used in that class. The science teacher (or teachers) of the students in that class will be asked to complete a science teacher questionnaire, which in many respects will parallel the mathematics