

Chapter 2



Developing the TIMSS 2007 Mathematics and Science Assessments and Scoring Guides

Graham J. Ruddock, Christine Y. O'Sullivan, Alka Arora, and Ebru Erberber

2.1 Overview

The mathematics and science assessments for TIMSS 2007 were developed over a period of 2 years, from January 2005 to November 2006. The process incorporated the expertise of mathematics and science educators and test development specialists from all over the world. In particular, the TIMSS & PIRLS International Study Center worked with the Science and Mathematics Item Review Committee (SMIRC), an international committee of prominent mathematics and science experts who were nominated by participating countries and represented a range of nations and cultures.¹ Their responsibilities were threefold: to review and revise items, check for mathematical and scientific accuracy, and make certain that the items fit the specifications contained within the *TIMSS 2007 Assessment Frameworks* (Mullis, Martin, Ruddock, O'Sullivan, Arora, & Erberber, 2005).

First, the mathematics and science assessment frameworks for TIMSS 2003 were reviewed and updates were made for TIMSS 2007. Because approximately half of the mathematics and science assessment items were released to the public following the publication of the TIMSS 2003 results, a large number of replacement items were newly developed for TIMSS 2007. Item writing was accomplished in large measure by the National Research Coordinators (NRCs), with support and training from the TIMSS & PIRLS

¹ See Appendix A for a list of members of the Science and Mathematics Item Review Committee.

International Study Center. Two task forces, one for mathematics and one for science,² were convened to assist in managing the resulting pool of items.

This chapter describes the test development process in more detail. An overview is shown in Exhibit 2.1.

Exhibit 2.1 Overview of the TIMSS 2007 Frameworks and Test Development Process

Date(s)		Group and Activity
January	2005	TIMSS & PIRLS International Study Center Review of TIMSS 2003 Assessment Frameworks Propose revisions for TIMSS 2007 Assessment Frameworks to be discussed in the First National Research Coordinators Meeting
February	2005	First National Research Coordinators Meeting (Cairo) Review of proposed changes to TIMSS 2003 Assessment Frameworks
March	2005	TIMSS & PIRLS International Study Center Compile TIMSS 2007 Content Domains and Percentages Survey to be completed by the National Research Coordinators Incorporating results from the survey and the First National Research Coordinators Meeting, prepare draft TIMSS 2007 Assessment Frameworks
April	2005	First Science and Mathematics Item Review Committee Meeting (London) Improve and review draft of TIMSS 2007 Assessment Frameworks
May	2005	Second National Research Coordinators Meeting (Amsterdam) Develop field test item pool using TIMSS Item-writing Guidelines
June-July	2005	TIMSS & PIRLS International Study Center Review, refine, and edit field test items in the pool Develop additional items to cover framework
August	2005	Science and Mathematics Task Forces (Boston) Review and revise field test items Develop additional items to cover framework
August	2005	Second Science and Mathematics Item Review Committee Meeting (Kyoto) Review field test item pool and select preferred and alternate items for field test to be reviewed in the Third National Research Coordinators Meeting
September	2005	TIMSS & PIRLS International Study Center Incorporating improvements and revisions from the Second Science and Mathematics Item Review Committee Meeting, finalize and publish the TIMSS 2007 Assessment Frameworks
November	2005	Third National Research Coordinator Meeting (Ljubljana) Review and approve preferred field test items
November	2005	TIMSS & PIRLS International Study Center Assemble field test item blocks

2 The mathematics task force consisted of Graham Ruddock, TIMSS Mathematics Coordinator; Robert Garden, former TIMSS Mathematics Coordinator; and Mary Lindquist, former president of National Council of Teachers of Mathematics. The science task force consisted of Christine O'Sullivan, TIMSS Science Coordinator; Audrey Champagne, State University of New York at Albany; and Jackie Heaton, University of Aberdeen.

Exhibit 2.1 Overview of the TIMSS 2007 Frameworks and Test Development Process (Continued)

Date(s)		Group and Activity
December	2005	TIMSS & PIRLS International Study Center Finalize field-test instruments and post them on the TIMSS & PIRLS International Study Center website for downloading
January	2006	TIMSS & PIRLS International Study Center Conduct a pilot test of constructed-response items Prepare draft scoring guides for constructed-response items
February	2006	Science and Mathematics Task Forces (Boston) Finalize scoring guides for constructed-response items Develop scoring training materials for the Fourth National Research Coordinators Meeting
March	2006	Fourth National Research Coordinators Meeting (Malta) Field test scoring training
March–April	2006	TIMSS 2007 field test administration
June	2006	TIMSS & PIRLS International Study Center Review field test item statistics Propose items for the main data collection to be discussed in the Third Science and Mathematics Item Review Committee Meeting
July	2006	Third Science and Mathematics Item Review Committee Meeting (St. Petersburg) Review proposed sets of items in the light of field test results
August	2006	Fifth National Research Coordinators Meeting (Riga) Review and approve proposed sets of items for the main data collection
August	2006	TIMSS & PIRLS International Study Center Finalize main survey materials and post them on the TIMSS & PIRLS International Study Center website for downloading
October	2006	TIMSS & PIRLS International Study Center Update scoring guides for constructed-response items Develop scoring training materials for the scoring training for the Southern Hemisphere countries
October	2006	Southern Hemisphere scoring training (Melbourne)
November	2006	TIMSS & PIRLS International Study Center Incorporate revisions from the Southern Hemisphere scoring training, finalize scoring guides for constructed-response items and post them on the TIMSS & PIRLS International Study Center website for downloading

2.2 Updating the TIMSS 2007 Assessment Frameworks

The *TIMSS 2007 Assessment Frameworks* closely resembles its corresponding publication for TIMSS 2003. However, there were three important areas of updates in the assessment for TIMSS 2007.

- In the Mathematics and Science Frameworks, the content domains were presented separately for the fourth and eighth grades, and there was a concerted effort to better reflect fourth-grade curricula. At both grades, there was an effort to consolidate the major content areas and adjust the assessment topic areas and objectives to make them more appropriate and feasible in the context of a large-scale international assessment.
- To increase the potential for analyzing and reporting the mathematics and science results according to cognitive domains, the U.S. National Center for Education Statistics provided support to examine and refine the mathematics cognitive domains used in TIMSS 2003 and report the results.³ Accordingly, the cognitive domains in the Mathematics and Science Frameworks were updated to enable reporting by cognitive domains in the TIMSS 2007 International Reports.
- The assessment design was updated to increase students' response time in TIMSS 2007 because TIMSS 2003 had been somewhat speeded in some countries. The allocation of blocks to booklets was modified to include fewer blocks in a booklet and to have the design fully balanced. That is, each booklet in TIMSS 2007 included 2 mathematics blocks and 2 science blocks, with half the booklets having the mathematics blocks first and half having the science blocks first. The time provided to students to complete a block was increased—from 12 to 18 minutes at fourth grade and from 15 to 22.5 minutes at eighth grade.

In preparation for the first meeting of the TIMSS 2007 National Research Coordinators, the TIMSS & PIRLS International Study Center staff met with the Mathematics and Science Coordinators to discuss the updates to Mathematics and Science Frameworks and with consultants from Educational Testing Service to discuss the updates to the assessment design.

The proposed updates were discussed during the first NRC meeting in February 2005. The NRCs were pleased with the updated, simplified

3 For details of this cognitive study, see Mullis, I.V.S., Martin, M.O., & Foy, P. (2005). *IEA's TIMSS 2003 international report on achievement in the mathematics cognitive domains*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

design that provided students with increased response time, even though this meant countries measuring trend would be required to participate in a bridge study. In addition to full participation in TIMSS 2007, trend countries would need to administer a subset of TIMSS 2003 booklets in TIMSS 2007 under the TIMSS 2003 conditions. The NRCs also were supportive of the recommendations for updating the content and cognitive domains in the Mathematics and Science Frameworks, and were appreciative of the efforts to report results by the cognitive domains. They recommended conducting a survey to gather further information about how to update the content and cognitive domains.

Following the NRC meeting, TIMSS & PIRLS International Study Center staff distributed the survey questionnaires and asked NRCs to indicate their country's preferences with regard to the content domains (algebra, geometry, measurement, physics, chemistry, etc.) that would characterize mathematics and science in the assessment and the specific topics in each domain that would be addressed by the assessment objectives. The responses to the questionnaires from participating countries showed broad support for the proposed changes to the assessment frameworks and provided valuable guidance in drafting the 2007 frameworks in March 2005. The draft frameworks were reviewed by members of the SMIRC in April 2005, and the final *TIMSS 2007 Assessment Frameworks* were published in September 2005.

For the TIMSS 2007 frameworks, a decision was made to separate the content domains for both mathematics and science by grade because of the increasing complexity of the subject matter and the introduction, at the eighth grade, of topics not covered at the fourth grade. This also allowed the fourth grade content domains to be renamed, where appropriate, to better reflect the content being assessed. The cognitive domains were streamlined based on information stemming from the study of the skills and abilities used to answer mathematics items in the 2003 assessment.

The major revisions in the mathematics content domains were organizational. At fourth grade, the previous number and algebra domains were combined into a revised number domain, while the previous measurement and geometry domains were restructured into a new geometric shapes and measures domain. The 2003 data domain was renamed “data display”, better reflecting what was being assessed at this grade. At the topic

level, the fourth grade topics were streamlined, and topics where minimal content was assessed were combined with others.

At eighth grade, the 2003 measurement domain was eliminated, and the topics covered were redistributed to geometry (length, area, volume, angle, perimeter, and circumference) or number (time, speed, mass/weight, and temperature).

The cognitive domains for mathematics were reduced from four to three: knowing, applying, and reasoning. The previous two cognitive domains, using concepts and solving routine problems, were split across the three new domains. This change allowed the cognitive structure of the assessment frameworks to be the same for mathematics and science.

The revisions made to the science assessment framework were mostly of an organizational nature. Since the content domains for the fourth and eighth grades were presented separately, the domains of chemistry and physics at the fourth grade were combined into physical science. In addition, at the fourth grade, the topics covered in the environmental science domain were moved to life science and earth science. In TIMSS 2003, environmental science items had been included in these domains, because environmental science was not a reporting strand at the fourth grade. Although the environmental science domain was a reporting strand in 2003 at eighth grade, this strand was eliminated for TIMSS 2007 and its topics moved to biology and earth science, where a new topic area was added—Earth’s resources, their use, and conservation. While topic areas at both grades were combined, the content remained essentially the same, with the exception of the topic of sound at the fourth grade. In the 2003 assessment, this topic area was not assessed, however, committee members decided that fourth grade students had a rudimentary knowledge of sound, and this area now should be assessed under physical science.

The topics that made up the science cognitive domains remained essentially the same as those contained within the TIMSS 2003 frameworks. However, the domain names did change so that they would be identical to those of mathematics. Thus, factual knowledge, conceptual understanding, and reasoning and analysis became knowing, applying, and reasoning, respectively. In addition, some of the topics included within the domains were moved from one cognitive domain to another based on current thinking about cognitive processes.

2.3 Mathematics Assessment Framework

The TIMSS 2007 mathematics framework is presented in full in the *TIMSS 2007 Assessment Frameworks*. The basic structure of the mathematics framework is defined by two dimensions, a content and a cognitive dimension, which remains unchanged from TIMSS 2003 and is illustrated in Exhibit 2.2. Exhibit 2.3 shows the topic areas within the content domains.

Exhibit 2.2 Target Percentages of the TIMSS 2007 Mathematics Assessment Devoted to Content and Cognitive Domains by Grade Level

Fourth Grade Content Domains		Percentages	
Number		50%	
Geometric Shapes and Measures		35%	
Data Display		15%	
Eighth Grade Content Domains		Percentages	
Number		30%	
Algebra		30%	
Geometry		20%	
Data and Chance		20%	
Cognitive Domains		Percentages	
		Fourth Grade	Eighth Grade
Knowing		40%	35%
Applying		40%	40%
Reasoning		20%	25%

Exhibit 2.3 Topic Areas Included in the Mathematics Content Domains by Grade Level

Fourth Grade Content Domains	Fourth Grade Topic Areas
Number	Whole numbers Fractions and decimals Number sentences Patterns and relationships
Geometric Shapes and Measures	Lines and angles Two- and three-dimensional shapes Location and movement
Data Display	Reading and interpreting Organizing and representing
Eighth Grade Content Domains	Eighth Grade Topic Areas
Number	Whole numbers Fractions and decimals Integers Ratio, proportion, and percent
Algebra	Patterns Algebraic Expressions Equations/formulas and functions
Geometry	Geometric shapes Geometric measurement Location and movement
Data and Chance	Data organization and representation Data interpretation Chance

2.4 Science Assessment Framework

As in mathematics, the science assessment framework is fully detailed in the *TIMSS 2007 Assessment Frameworks*. It is organized along two dimensions: content and cognitive. The content dimension at the fourth grade is made up of three domains: life science, physical science, and earth science. The four content domains at the eighth grade are: biology, chemistry, physics, and earth science. There are three cognitive domains at both fourth and eighth grades: knowing, applying, and reasoning. The target percentages for the content and cognitive domains are shown in Exhibit 2.4 and the topic areas are listed in Exhibit 2.5.

Exhibit 2.4 Target Percentages of the TIMSS 2007 Science Assessment Devoted to Content and Cognitive Domains by Grade Level

Fourth Grade Content Domains		Percentages	
Life Science		45%	
Physical Science		35%	
Earth Science		20%	
Eighth Grade Content Domains		Percentages	
Biology		35%	
Chemistry		20%	
Physics		25%	
Earth Science		20%	
Cognitive Domains		Percentages	
		Fourth Grade	Eighth Grade
Knowing		40%	30%
Applying		35%	35%
Reasoning		25%	35%

Exhibit 2.5 Topic Areas Included in the Science Content Domains by Grade Level

Fourth Grade Content Domains	Fourth Grade Topic Areas
Life Science	<ul style="list-style-type: none"> Characteristics and life processes of living things Life cycles, reproduction, and heredity Interaction with the environment Ecosystems Human health
Physical Science	<ul style="list-style-type: none"> Classification and properties of matter Physical states and changes in matter Energy sources, heat, and temperature Light and sound Electricity and magnetism Forces and motion
Earth Science	<ul style="list-style-type: none"> Earth's structure, physical characteristics, and resources Earth's processes, cycles, and history Earth in the solar system
Eighth Grade Content Domains	Eighth Grade Topic Areas
Biology	<ul style="list-style-type: none"> Characteristics, classification, and life processes of organisms Cells and their functions Life cycles, reproduction, and heredity Diversity, adaptation, and natural selection Ecosystems Human health
Chemistry	<ul style="list-style-type: none"> Classification and composition of matter Properties of matter Chemical change
Physics	<ul style="list-style-type: none"> Physical states and changes in matter Energy transformations, heat, and temperature Light Sound Electricity and magnetism Forces and motion
Earth Science	<ul style="list-style-type: none"> Earth's structure and physical features Earth's processes, cycles, and history Earth's resources, their use, and conservation Earth in the solar system and the universe

2.5 Developing Mathematics and Science Items and Scoring Guides

Because approximately half of the fourth- and eighth-grade items from TIMSS 2003 were kept secure to be readministered in the 2007 assessment, blueprints for mathematics and science item development were developed to ensure that the newly developed items met the guidelines laid out in the 2007 assessment framework. The blueprints were created by:

- Estimating the number of items needed in the assessment based on the total score points and percentage of score points in each content domain specified in the framework
- Distributing this number of items across the mathematics and science main topic areas according to their breadth of content
- Accounting for the number of trend items already included in each topic area
- Ensuring coverage of the cognitive domains and the appropriate numbers of multiple-choice and constructed-response items (the frameworks specify approximately 50% each)
- Scaling up the number of items to be developed to allow for attrition during the item selection and field-testing process.

This section describes the test development process and includes a consideration of trend items, development of the international item pool, item review and revision, problem-solving sets, field testing, item selection for the main data collection, development of scoring guides for constructed-response items, and scoring training.

2.5.1 Trend Items

The mathematics and science trend items from 2003 were mapped into the content and cognitive categories described in the TIMSS 2007 frameworks. The results are shown in Exhibits 2.6 and 2.7.

In mathematics at fourth grade, the number of multiple-choice and constructed-response items was about the same in each domain, so the newly developed items needed to maintain this balance. This was not the situation at eighth grade, where multiple-choice items predominated, except in the data and chance domain. For this grade, more new constructed-response items were needed to redress the balance.

Science included both multiple-choice and constructed-response items in each domain. At fourth grade, the number of multiple-choice and

constructed-response trend items was about the same. Eighth grade had more multiple-choice items, hence, a larger proportion of constructed-response items needed to be developed for eighth grade.

Exhibit 2.6 Mathematics Trend Items by Grades, Content and Cognitive Domains, and Item Format

Content Domain	Fourth Grade Trend Items			Eighth Grade Trend Items		
	Multiple Choice	Constructed Response	Total	Multiple Choice	Constructed Response	Total
Number	24	27	51	18	10	28
Algebra*	–	–	–	14	8	22
Geometry**	12	9	21	18	9	27
Data and Chance***	5	4	9	6	12	18
Total	41	40	81	56	39	95

Cognitive Domain	Multiple Choice	Constructed Response	Total	Multiple Choice	Constructed Response	Total
Knowing	16	7	23	23	8	31
Applying	18	18	36	26	19	45
Reasoning	7	15	22	7	12	19
Total	41	40	81	56	39	95

* Prealgebraic concepts are included in the Number content domain at the fourth grade.

** Called Geometric Shapes and Measures at the fourth grade

*** Called Data Display at the fourth grade.

Exhibit 2.7 Science Trend Items by Grades, Content and Cognitive Domains, and Item Format

Content Domain	Fourth Grade Trend Items			Eighth Grade Trend Items		
	Multiple Choice	Constructed Response	Total	Multiple Choice	Constructed Response	Total
Biology*	22	10	32	16	18	34
Physical Science	13	16	29	–	–	–
Chemistry	–	–	–	9	5	14
Physics	–	–	–	16	7	23
Earth Science	5	10	15	10	13	23
Total	40	36	76	51	43	94

Cognitive Domain	Multiple Choice	Constructed Response	Total	Multiple Choice	Constructed Response	Total
Knowing	19	13	32	30	9	39
Applying	15	10	25	15	18	33
Reasoning	6	13	19	6	16	22
Total	40	36	76	51	43	94

* Called Life Science at the fourth grade.

2.5.2 Developing the International Item Pool for TIMSS 2007

During the second NRC meeting in May 2005, participants from 60 countries attended the item-writing workshop for TIMSS 2007. The TIMSS & PIRLS International Study Center provided specific instructions on how to write multiple-choice and constructed-response items in accordance with an item-writing manual that had been developed for TIMSS 2007. Participants then were organized into four subgroups—fourth grade mathematics, eighth grade mathematics, grade fourth science, and eighth grade science. These subgroups were further subdivided into content domain groups. The item-writing effort was very successful, yielding a large number of draft items—for mathematics, approximately 210 and 230 items at fourth and eighth grades, respectively, and for science, approximately 200 and 120 items at these respective grades. Also, several countries sent additional items to the TIMSS & PIRLS International Study Center in the weeks that followed.

2.5.3 Item Review and Revision

Following item development, members of the task force reviewed and revised the items and wrote additional items to cover specific areas of the frameworks that had not been addressed. The resulting item pool consisted of 997 items, of which 472 were in mathematics and 525 were in science. The items then were reviewed by the SMIRC. For this review, members of the committee formed two subgroups, a mathematics group and a science group. Members of the subgroups reviewed items for content accuracy, grade appropriateness, and framework fit.

To increase efficiency, the field test blocks were organized to represent the desired assessment as much as possible. So far as the field test was successful, then materials did not have to be reformatted. The replacement items were prepared in parallel blocks, so they could be incorporated into the assessment blocks with minimal disruption. Thus, the field-test item pool was divided into two sets, “preferred” and “alternate”, and the items then organized into two sets of item blocks. One set of item blocks contained the “preferred” items and the second set contained the “alternate” items. These blocks of items were reviewed at the third NRC meeting that was held in Slovenia in November 2005. NRCs made suggestions for revising certain items, mostly based on concerns about translation issues. They also raised concerns about the grade appropriateness of a few items, and these were subsequently eliminated.

2.5.4 Item Sets

Several extended reasoning tasks or item sets of related problems were developed by the SMIRC and by members of the task force for inclusion in TIMSS 2007. These were written to specifically measure aspects of the frameworks that were difficult to assess using discrete items and, generally, were classified as reasoning. For these tasks, the number of possible points, typically 3 to 6 points, depended on the requirements for students to successfully complete the task.

In mathematics, the extended tasks involved patterns and their generalization with scenarios based on real-life contexts. Two were included in the final assessment for fourth grade and four for eighth grade.

In science, the extended tasks addressed aspects of science inquiry based on experimental set-ups or student investigations. They were reviewed and revised at each SMIRC and task force meeting. Among the tasks that were field tested, two were included at the fourth grade level and five at the eighth grade level in the final assessment.

2.5.5 Field Test

The newly developed fourth- and eighth-grade items were field tested internationally from March to April 2006. In total, 31 countries participated in the fourth grade field test and 45 countries participated in the eighth grade field test. The field test in each country was administered to a random sample of a minimum of 25 schools. Approximately twice the number of items were field tested than were needed for the TIMSS 2007 assessment. A total of 350 items were included in the fourth grade field test, 192 in mathematics and 158 in science. At the eighth grade, a total of 415 items were included in the field test, 214 in mathematics and 201 in science. Since some constructed-response items contributed 2 score points, this corresponded to a total of 203 score points in mathematics and 180 in science at the fourth grade, and 283 score points in mathematics and 240 in science at the eighth grade.

2.5.6 Item Selection for the TIMSS 2007 Data Collection

The selection of items for the TIMSS 2007 data collection was based on an item analysis of the international results of the field test. Data almanacs containing basic item statistics for each country and internationally were produced, including the following:

- Difficulty levels for each item
- How well items discriminated between high- and low-performing students
- The effectiveness of distractors in multiple-choice items
- The frequency of occurrence of diagnostic codes used in the scoring guides
- Scoring reliability for constructed-response items.

The TIMSS & PIRLS International Study Center and the mathematics and science coordinators reviewed data from the field test in June 2006. Items were selected that discriminated well, had a range of difficulties, and covered the cognitive and content domains, and draft blocks of these items were assembled, including a pool of alternate items.

The SMIRC reviewed these draft blocks in July 2006. Some items were revised slightly and other replaced with items from the alternate pool. The modified draft blocks then were reviewed by the NRCs at the fifth NRC meeting held in August 2006. The NRCs recommended some further changes to items and asked that a few items be replaced. A total of 196 new items at the fourth grade and 240 items at the eighth grade were approved by the NRCs for inclusion in the TIMSS 2007 data collection. The final assessments were made up of 353 items at fourth grade and 429 items at eighth grade, including both trend and new items.

Exhibits 2.8 and 2.9 show the distribution of new and trend items in the TIMSS 2007 mathematics and science assessments by item format for fourth and eighth grades, respectively. They reflect the number of individual items and all item subparts included in multipart items.

The percentage of score points contributed to the assessments by constructed-response items for both mathematics and science and at both grade levels ranged from 50 to 55 percent.

Exhibit 2.8 Distribution of New and Trend Items in the TIMSS 2007 by Subject and Item Format – Fourth Grade

Item Format	Number of Items				
	New Items	Trend Items	Total (New+Trend)	Total Score Points	Percentage of Score Points
Mathematics Items					
Multiple Choice	55	41	96	96	50%
Constructed Response	43	40	83	96	50%
Total Mathematics Items	98	81	179	192	
Science Items					
Multiple Choice	53	40	93	93	48%
Constructed Response	45	36	81	101	52%
Total Science Items	98	76	174	194	
All Items					
Multiple Choice	108	81	189	189	49%
Constructed Response	88	76	164	197	51%
Total Items	196	157	353	386	

Exhibit 2.9 Distribution of New and Trend Items in the TIMSS 2007 by Subject and Item Format – Eighth Grade

Item Format	Number of Items				
	New Items	Trend Items	Total (New+Trend)	Total Score Points	Percentage of Score Points
Mathematics Items					
Multiple Choice	61	56	117	117	49%
Constructed Response	59	39	98	121	51%
Total Mathematics Items	120	95	215	238	
Science Items					
Multiple Choice	56	51	107	107	45%
Constructed Response	64	43	107	133	55%
Total Science Items	120	94	214	240	
All Items					
Multiple Choice	117	107	224	224	47%
Constructed Response	123	82	205	254	53%
Total Items	240	189	429	478	

2.5.7 Developing Scoring Guides for Constructed-response Items

Constructed-response items made up approximately half of the total assessment time in the TIMSS 2007 test, with each of these items having its own scoring guide.

2.5.7.1 The TIMSS General Scoring Method

TIMSS 2007, as in previous TIMSS assessments, included constructed-response items and used the same approach to scoring. Constructed-response questions generally are worth 1 or 2 score points, depending on the nature of the task or skills required to complete it. Typically, constructed-response items worth 1 score point require a numerical response in mathematics or a brief descriptive response in science, while those worth 2 score points require students to show their work or provide explanations using words and/or diagrams to demonstrate their conceptual understanding.

The generalized mathematics and science scoring guidelines that were developed and applied for TIMSS 2003 also were used in TIMSS 2007. The scoring guidelines are shown in Exhibit 2.10.

**Exhibit 2.10 TIMSS Generalized Scoring Guidelines for Mathematics and Science
Constructed-response Items**

Score Points for 1-point Items	
Mathematics	Science
<p>1 Point</p> <p>A 1-point response is correct. The response indicates that the student has completed the task correctly.</p>	<p>1 Point</p> <p>A 1-point response is correct. The response indicates that the student has completed the task correctly.</p>
<p>0 Points</p> <p>A 0-point response is completely incorrect, irrelevant, or incoherent.</p>	<p>0 Points</p> <p>A 0-point response is completely incorrect, irrelevant, or incoherent.</p>
Score Points for 2-point Items	
<p>2 Points</p> <p>A 2-point response is complete and correct. The response demonstrates a thorough understanding of the mathematical concepts and/or procedures embodied in the task.</p> <ul style="list-style-type: none"> Indicates that the student has completed the task, showing mathematically sound procedures Contains clear, complete explanations and/or adequate work when required 	<p>2 Points</p> <p>A 2-point response is complete and correct. The response demonstrates a thorough understanding of the science concepts and/or procedures embodied in the task.</p> <ul style="list-style-type: none"> Indicates that the student has completed all aspects of the task, showing the correct application of scientific concepts and/or procedures Contains clear, complete explanations when required
<p>1 Point</p> <p>A 1-point response is only partially correct. The response demonstrates only a partial understanding of the mathematical concepts and/or procedures embodied in the task.</p> <ul style="list-style-type: none"> Addresses some elements of the task correctly but may be incomplete or contain some procedural or conceptual flaws May contain a correct solution with incorrect, unrelated, or no work and/or explanation when required May contain an incorrect solution but applies a mathematically appropriate process 	<p>1 Point</p> <p>A 1-point response is only partially correct. The response demonstrates only a partial understanding of the science concepts and/or procedures embodied in the task.</p> <ul style="list-style-type: none"> Addresses some elements of the task correctly but may be incomplete or contain some procedural or conceptual flaws May contain a correct answer but with an incomplete explanation when required May contain an incorrect answer but with an explanation indicating a correct understanding of some of the scientific concepts
<p>0 Points</p> <p>A 0-point response is completely incorrect, irrelevant, or incoherent.</p>	<p>0 Points</p> <p>A 0-point response is seriously inaccurate or inadequate, irrelevant, or incoherent.</p>

Each constructed-response item has its own scoring guide that utilizes a two-digit scoring scheme to provide diagnostic information. The first digit designates the correctness level of the response: 2 for a 2-point response, 1 for a 1-point response, and 0 for an incorrect response. The second digit, combined with the first, represents a diagnostic code used to identify specific

types of approaches, strategies, or common errors and misconceptions. A second digit of 0 to 5 may be used for a predefined international code at each correctness level, while a second digit of 9 corresponds to “other” types of responses that fall within the appropriate correctness level but do not fit any of the predefined international codes. A special code (99) is given for completely blank responses. In general, only a few diagnostic codes are used to track high frequency correct or partial approaches or common misconceptions and errors. In addition to the international codes, second digit codes of 7 and 8 may be used by national centers to monitor specific responses not already captured by the internationally defined codes. The general TIMSS two-digit scoring scheme is summarized in Exhibit 2.11.

Exhibit 2.11 TIMSS Two-digit Scoring Scheme for Constructed-response Items

2-Point Items		1-Point Items	
Correctness Level	International Code(s)	Correctness Level	International Code(s)
Correct Responses	20–25: Category/method #1–#5 29: Other correct method	Correct Responses	10–15: Category/method #1–#5 19: Other correct method
Partial Responses	10–15: Category/method #1–#5 19: Other partial method	Incorrect Responses	70–75: Misconception/error #1–#5 79: Other error
Incorrect Responses	70–75: Misconception/error #1–#5 79: Other error	Blank	99
Blank	99		

2.5.7.2 Developing the TIMSS 2007 Scoring Guides

Scoring guides were written at the same time items were drafted. This helped ensure that the scoring guides captured what the items purported to measure and helped clarify the language of the items. The scoring guides were reviewed with the items by both the task force and members of the SMIRC. After the field test had been assembled, a pilot test of constructed-response items was conducted in English-speaking countries to collect student responses to use in developing scoring training materials for the field test. These responses helped refine the scoring guide, as well as clarify categories. They also helped in item revisions since ambiguities could be recognized by the way in which students responded to items. Selected student responses were included as examples in the scoring guides. Field test scoring training took place in March 2006 at the fourth NRC meeting. Scoring guides were further refined at this meeting.

A review of the field-test item statistics, however, showed that further refinements to the scoring guides could be made. For example, some of the

categories were not being used and some guides showed no discrimination between students who received full credit for an item and those who received partial credit. In addition to the data, feedback was received from the scoring trainers in the participating countries about their experiences with the scoring guides. All this information was considered when selecting items for TIMSS 2007.

Where necessary, scoring guides were revised for items chosen for the TIMSS 2007 data collection. They were reviewed again by the SMIRC and NRCs during the final item review. Scoring guides for the trend items remained unchanged from the versions used in 2003.

2.5.7.3 Scoring Training Materials and Procedures

Scoring training for TIMSS 2007 was conducted in October 2006 for countries in the Southern Hemisphere and all of the scoring guides and training materials were finalized at this training. The training was replicated in March 2007 for countries in the Northern Hemisphere. The training materials used and the procedures followed for scoring training were very similar to those for previous TIMSS surveys. Representatives from countries participating in the survey were given information about the TIMSS scoring method and then trained using a selection of items chosen to illustrate the various types of constructed-response items. Materials for the scoring training were posted on the TIMSS & PIRLS International Study Center website prior to the meeting, and participants brought the materials with them. The training materials for each item included the scoring guide, a set of student papers illustrating the different score levels, and a set of practice papers.

The purpose of the scoring training was twofold: to present a model for use within each participating country and to give participants an opportunity to practice and resolve scoring issues with the most difficult items.

The following general procedures were followed while training each item:

- Participants read the item and its scoring guide.
- Trainers discussed the rationale and methodology of the scoring guide.
- Trainers presented and discussed the set of prescored example student responses.
- Participants scored the set of practice student responses.

2.6 Assessment Booklet Design

The TIMSS design for 2007 divided the 353 items at fourth grade and 429 items at eighth grade into 28 item blocks at each grade, 14 mathematics blocks labeled M01 through M14, and 14 science blocks labeled S01 through S14. Each block contained either mathematics items only or science items only. This general block design is shown in Exhibit 2.12 and is the same for each grade level. However, the assessment time was 18 minutes for fourth grade blocks and 22.5 minutes for eighth grade blocks. At fourth and eighth grades, seven blocks (the odd-numbered ones) contained secure items from 2003 that were used to measure trends, and seven blocks (the even-numbered ones) contained items newly developed for TIMSS 2007.

Exhibit 2.12 General Design of the TIMSS Matrix-sampling Blocks

Mathematics Blocks	Source of Items	Science Blocks	Source of Items
M01	Block M05 from TIMSS 2003	S01	Block S14 from TIMSS 2003
M02	New items for TIMSS 2007	S02	New items for TIMSS 2007
M03	Block M06 from TIMSS 2003	S03	Block S05 from TIMSS 2003
M04	New items for TIMSS 2007	S04	New items for TIMSS 2007
M05	Block M07 from TIMSS 2003	S05	Block S06 from TIMSS 2003
M06	New items for TIMSS 2007	S06	New items for TIMSS 2007
M07	Block M08 from TIMSS 2003	S07	Block S07 from TIMSS 2003
M08	New items for TIMSS 2007	S08	New items for TIMSS 2007
M09	Block M11 from TIMSS 2003	S09	Block S08 from TIMSS 2003
M10	New items for TIMSS 2007	S10	New items for TIMSS 2007
M11	Block M12 from TIMSS 2003	S11	Block S11 from TIMSS 2003
M12	New items for TIMSS 2007	S12	New items for TIMSS 2007
M13	Block M14 from TIMSS 2003	S13	Block S12 from TIMSS 2003
M14	New items for TIMSS 2007	S14	New items for TIMSS 2007

In the TIMSS 2007 design, the 28 blocks of items were distributed across 14 student booklets, as shown in Exhibit 2.13. Each booklet consisted of four blocks of items. To enable linking between booklets, each block appears in two booklets. The assessment time for individual students was 72 minutes at fourth grade and 90 minutes at eighth grade, which is comparable to that in the 1995, 1999, and 2003 assessments. The booklets were organized into 2 two-block sessions (Parts I and II), with a break in between each part.

Exhibit 2.13 Booklet Design for TIMSS 2007 – Fourth Grade and Eighth Grade

Student Achievement Booklet	Assessment Blocks			
	Part 1		Part 2	
Booklet 1	M01	M02	S01	S02
Booklet 2	S02	S03	M02	M03
Booklet 3	M03	M04	S03	S04
Booklet 4	S04	S05	M04	M05
Booklet 5	M05	M06	S05	S06
Booklet 6	S06	S07	M06	M07
Booklet 7	M07	M08	S07	S08
Booklet 8	S08	S09	M08	M09
Booklet 9	M09	M10	S09	S10
Booklet 10	S10	S11	M10	M11
Booklet 11	M11	M12	S11	S12
Booklet 12	S12	S13	M12	M13
Booklet 13	M13	M14	S13	S14
Booklet 14	S14	S01	M14	M01

2.6.1 Assembling Item Blocks

The assessment blocks were assembled to create a balance across blocks and booklets with respect to content domain, cognitive domain, and item format. Depending on the exact number of multiple-choice and constructed-response items in each block, the total number of mathematics items in a block ranged from 10–14 at fourth grade and 11–18 at eighth grade. In science, depending on the exact number of multiple-choice and constructed-response items in each block, the total number of science items in a block ranged from 10–14 at fourth grade and 12–18 at eighth grade.

2.6.2 Incorporating Trend Items

At fourth grade, 14 blocks of items from TIMSS 2003 were used in TIMSS 2007—7 from mathematics and 7 from science. These were renumbered as shown in Exhibit 2.14.

At eighth grade in mathematics, one block of items from 1999 and one block containing items from 1999 and 2003 were used in 2007 (see Exhibit 2.15). These are labeled M01 and M03, respectively. The remaining five trend blocks were first administered in 2003. At eighth grade in science, similar to mathematics, one block containing items from 1999 and one block containing items from 1999 and 2003 were used in 2007 and labeled S03 and S05, respectively. The remaining five trend blocks were first administered in 2003.

Exhibit 2.14 TIMSS 2007 Mathematics and Science Blocks – Fourth Grade: Number of Items from Trend Blocks and Score Points by Assessment Year

Block	Number of Items from Trend Blocks*	Score Points by Assessment Year		
		2003	2007	Total
Mathematics Blocks				
M01	M05(11)	12	0	12
M03	M06(12)	12	0	12
M05	M07(11)	12	0	12
M07	M08(12)	12	0	12
M09	M11(12)	14	0	14
M11	M12(13)	13	0	13
M13	M14(10)	11	0	11
M02, M04, M06, M08, M10, M12, M14	–	0	106	106
Mathematics Total	81	86	106	192
Science Blocks				
S01	S14(11)	12	0	12
S03	S05(11)	12	0	12
S05	S06(10)	11	0	11
S07	S07(11)	12	0	12
S09	S08(11)	13	0	13
S11	S11(11)	12	0	12
S13	S12(11)	13	0	13
S02, S04, S06, S08, S10, S12, S14	–	0	109	109
Science Total	76	85	109	194
Overall Total	157	171	215	386

* The number of items from each trend block is indicated in parentheses.

Exhibit 2.15 TIMSS 2007 Mathematics and Science Blocks – Eighth Grade: Number of Items from Trend Blocks and Score Points by Assessment Year

Block	Number of Items from Trend Blocks*	Score Points by Assessment Year			
		1999	2003	2007	Total
Mathematics Blocks					
M01	M05(13)	16	0	0	16
M03	M06(15)	8	7	0	15
M05	M07(12)	0	17	0	17
M07	M08(15)	0	15	0	15
M09	M11(14)	0	15	0	15
M11	M12(15)	0	16	0	16
M13	M14(11)	0	15	0	15
M02, M04, M06, M08, M10, M12, M14	–	0	0	129	129
Mathematics Total	95	24	85	129	238
Science Blocks					
S01	S14(12)	0	15	0	15
S03	S05(14)	15	0	0	15
S05	S06(15)	8	7	0	15
S07	S07(12)	0	15	0	14
S09	S08(14)	0	16	0	15
S11	S11(13)	0	15	0	15
S13	S12(14)	0	16	0	16
S02, S04, S06, S08, S10, S12, S14	–	0	0	133	133
Science Total	94	23	84	133	240
Overall Total	189	47	169	262	478

*The number of items from each trend block is indicated in parentheses.

2.6.3 Alignment with the Mathematics and Science Frameworks

The test development process for TIMSS 2007 resulted in fourth- and eighth-grade assessments that are aligned with the *TIMSS 2007 Assessment Frameworks*. Details of the coverage of the frameworks are presented in the following subsections, for each grade level.

2.6.3.1 Fourth Grade Assessment

Exhibit 2.16 shows the distribution of score points for the fourth grade mathematics assessment by content and cognitive domains. (The mathematics framework target percentages can be seen in Exhibit 2.2.) The content domains were all within 1 percent of the target percentages. The proportion of reasoning items was slightly higher than the target because of the distribution of trend items.

Exhibit 2.16 Distribution of Score Points in the TIMSS 2007 Mathematics Assessment by Content and Cognitive Domains – Fourth Grade

Content Domain	Cognitive Domain			Total Score Points	Percentage of Score Points
	Knowing	Applying	Reasoning		
Number	41	33	24	98	51%
Geometric Shapes and Measures	26	28	11	65	34%
Data Display	6	14	9	29	15%
Total Score Points	73	75	44	192	
Percentage of Score Points	38%	39%	23%		

Exhibit 2.17 shows the score point distribution for the fourth grade science assessment by content and cognitive domain. (The science framework target percentages can be seen in Exhibit 2.4.) The percentages for the content domains were met. The percentage of items assessing knowing was higher than the target percentage, and the percentage of items assessing reasoning was lower than the target percentage. This was due in part to some of the items being reclassified based on the movement of several topics from one cognitive domain to another. About 17 percent of the score points in science measured scientific inquiry, and these covered topics from all three content domains.

Exhibit 2.17 Distribution of Score Points in the TIMSS 2007 Science Assessment by Content and Cognitive Domains and Scientific Inquiry Strand – Fourth Grade

Content Domain	Cognitive Domain			Total Score Points	Percentage of Score Points	Scientific Inquiry Score Points
	Knowing	Applying	Reasoning			
Life Science	41	31	13	85	44%	3
Physical Science	24	23	20	67	35%	24
Earth Science	24	14	4	42	21%	6
Total Score Points	89	68	37	194		33
Percentage of Score Points	46%	35%	19%			17%

As specified in the *TIMSS 2007 Assessment Frameworks*, two item types were included in the survey—multiple-choice and constructed-response. Exhibit 2.18 shows the distribution of the fourth grade mathematics and science items by item type and content domain.

Exhibit 2.18 Number of Mathematics and Science Items in TIMSS 2007 by Item Type and Content Domain – Fourth Grade

Content Domain	Multiple Choice	Constructed Response	Total Number of Items
Mathematics Items			
Number	50	43	93
Geometric Shapes and Measures	32	28	60
Data Display	14	12	26
Total Mathematics Items	96	83	179
Science Items			
Life Science	42	32	74
Physical Science	35	29	64
Earth Science	16	20	36
Total Science Items	93	81	174

TIMSS reports trends in student achievement in mathematics and science in the content domains of each subject. To facilitate linking to previous assessments, TIMSS 2007 at fourth grade included items from 2003 (TIMSS was not conducted at fourth grade in 1999). The number of score points in mathematics and science contributed by items used in TIMSS 2003 fourth grade can be seen in Exhibit 2.19.

Exhibit 2.19 Number of Score Points in TIMSS 2007 from Each Assessment Year by Mathematics and Science Content Domains – Fourth Grade

Content Domain	From 2003	New in 2007	Total 2007
Mathematics			
Number	55	43	98
Geometric Shapes and Measures	21	44	65
Data Display	10	19	29
Total in Mathematics	86	106	192
Science			
Life Science	35	50	85
Physical Science	32	35	67
Earth Science	18	24	42
Total in Science	85	109	194

The block and booklet design for 2007 ensured that the student booklets contained an appropriate balance of mathematics and science content. Exhibit 2.20 shows the number of mathematics and science score points in each fourth grade booklet.

Exhibit 2.20 Number of Score Points in TIMSS 2007 in Each Booklet by Mathematics and Science Content Domains – Fourth Grade

Content Domain	Booklet													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Mathematics														
Number	15	12	11	12	13	13	14	16	17	16	14	15	14	14
Geometric Shapes and Measures	7	11	11	8	10	12	10	9	9	10	8	6	10	9
Data Display	4	3	5	7	5	3	4	4	4	3	5	4	2	4
Total in Mathematics	26	26	27	27	28	28	28	29	30	29	27	25	26	27
Science														
Life Science	10	11	12	13	15	14	14	14	10	9	11	13	13	11
Physical Science	13	10	8	7	7	7	8	10	11	11	10	10	10	12
Earth Science	4	6	8	7	5	7	7	6	6	6	7	6	5	4
Total in Science	27	27	28	27	27	28	29	30	27	26	28	29	28	27

2.6.3.2 Eighth Grade Assessment

Exhibit 2.21 shows the distribution of score points for the eighth grade mathematics assessment by content and cognitive domain. For both content and cognitive domains, the percentage of score points was within 1 percent of the target percentage (see Exhibit 2.2 for target percentages).

Exhibit 2.21 Distribution of Score Points in the TIMSS 2007 Mathematics Assessment by Content and Cognitive Domains – Eighth Grade

Content Domain	Cognitive Domain			Total Score Points	Percentage of Score Points
	Knowing	Applying	Reasoning		
Number	28	33	11	72	30%
Algebra	33	15	21	69	29%
Geometry	8	29	13	50	21%
Data and Chance	14	21	12	47	20%
Total Score Points	83	98	57	238	
Percentage of Score Points	35%	41%	24%		

Exhibit 2.22 shows the distribution of score points across content and cognitive domains in the eighth grade science assessment. The target percentages (see Exhibit 2.4) were met for the content domains, but the percentages for the cognitive domains were higher than the target percent for knowing and applying and lower for reasoning. This was similar to fourth grade and was due in part to reclassification of items based on changes in the cognitive topic areas. Items that measured scientific inquiry accounted for approximately 24 percent of the score points in science. These items covered all three science content domains.

Exhibit 2.22 Distribution of Score Points in the TIMSS 2007 Science Assessment by Content and Cognitive Domains, and Scientific Inquiry Strand – Eighth Grade

Content Domain	Cognitive Domain			Total Score Points	Percentage of Score Points	Scientific Inquiry Score Points
	Knowing	Applying	Reasoning			
Biology	35	31	23	89	37%	20
Chemistry	16	18	12	46	19%	18
Physics	14	32	13	59	25%	14
Earth Science	24	16	4	44	19%	4
Total Score Points	89	97	52	238		56
Percentage of Score Points	37%	41%	22%			24%

Exhibit 2.23 shows the number of multiple-choice and constructed-response science items by their subject area content domains in TIMSS 2007.

Exhibit 2.23 Number of Mathematics and Science Items in TIMSS 2007 by Item Type and Content Domain – Eighth Grade

Content Domain	Multiple Choice	Constructed Response	Total Number of Items
Mathematics Items			
Number	35	28	63
Algebra	34	30	64
Geometry	31	16	47
Data and Chance	17	24	41
Total Mathematics Items	117	98	215
Science Items			
Biology	36	40	76
Chemistry	21	21	42
Physics	31	24	55
Earth Science	19	22	41
Total Science Items	107	107	214

To study trends in eighth grade student mathematics and science achievement, TIMSS 2007 included items from the 1999 and 2003 TIMSS assessments. Exhibit 2.24 shows that approximately 10 percent of points for the 2007 assessment in both mathematics and science came from items first administered in 1999, and approximately 30 percent of points came from items first administered in 2003.

Exhibit 2.24 Number of Score Points in TIMSS 2007 from Each Assessment Year by Mathematics and Science Content Domains – Eighth Grade

Content Domain	From 1999	From 2003	New in 2007	Total 2007
Mathematics				
Number	12	19	41	72
Algebra	1	26	42	69
Geometry	8	22	20	50
Data and Chance	3	18	26	47
Total in Mathematics	24	85	129	238
Science				
Biology	6	36	47	89
Chemistry	4	11	31	46
Physics	9	15	35	59
Earth Science	4	22	20	46
Total in Science	23	84	133	240

The number of score points and the distribution of score points across the mathematics and science content domains for each booklet in the eighth-grade assessment is shown in Exhibit 2.25.

Exhibit 2.25 Number of Score Points in TIMSS 2007 in Each Booklet by Mathematics and Science Content Domain – Eighth Grade

Content Domain	Booklet													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Mathematics														
Number	13	14	17	11	8	9	9	8	7	5	6	8	12	17
Algebra	6	5	5	9	10	9	11	11	9	13	13	15	15	7
Geometry	9	7	7	6	7	7	7	9	9	7	7	5	4	8
Data and Chance	6	7	6	10	11	9	6	5	7	8	7	4	4	4
Total in Mathematics	34	33	35	36	36	34	33	33	32	33	33	32	35	36
Science														
Biology	14	11	13	12	10	11	11	14	14	12	12	12	15	17
Chemistry	6	7	6	7	7	7	6	5	8	8	7	7	6	5
Physics	6	10	11	8	8	8	9	8	8	9	7	9	10	7
Earth Science	7	5	5	8	8	7	8	8	7	7	7	6	4	5
Total in Science	33	33	35	35	33	33	34	35	37	36	33	34	35	34
Total Overall	67	66	70	71	69	67	67	68	69	69	66	66	70	70

2.6.4 Item Release Policy

TIMSS 2007 is the fourth assessment in a series of regular 4-year studies, providing trend data from 1995, 1999, and 2003. As in previous assessments, the design for TIMSS 2007 and beyond (2011, 2015, etc.) provides for retaining some of the items for the measurement of trends and releasing some items into the public domain. In TIMSS 2007, 6 of the 14 assessment blocks in each subject were released after the assessment. The released item blocks include the two blocks containing the items from 1999, two blocks containing trend items from 2003, and two blocks used for the first time in 2007.⁴ As item blocks are released, new items will be developed to take their place. Exhibits 2.26 and 2.27 show the number of secure and released items from the TIMSS 2007 assessment for fourth and eighth grades listed by content domain.

4 Because TIMSS did not assess fourth grade students in 1999, the TIMSS 2007 fourth grade released blocks comprise four blocks from 2003 and two from 2007.

Exhibit 2.26 Number of Items in each Mathematics and Science Content Domain by Release Status in TIMSS 2007 – Fourth Grade

Content Domain	Secure	Released	Total
Mathematics			
Number	55	38	93
Geometric Shapes and Measures	36	24	60
Data Display	14	12	26
Total Mathematics	105	74	179
Science			
Life Science	45	29	74
Physical Science	38	26	64
Earth Science	20	16	36
Total Science	103	71	174
Total Overall	208	145	353

Exhibit 2.27 Number of Items in each Mathematics and Science Content Domain by Release Status in TIMSS 2007 – Eighth Grade

Content Domain	Secure	Released	Total
Mathematics			
Number	31	32	63
Algebra	47	17	64
Geometry	25	22	47
Data and Chance	23	18	41
Total Mathematics	126	89	215
Science			
Biology	47	29	76
Chemistry	26	16	42
Physics	33	22	55
Earth Science	21	20	41
Total Science	127	87	214
Total Overall	253	176	429

References

- Mullis, I.V.S., Martin, M.O., & Foy, P. (2005). *IEA's TIMSS 2003 international report on achievement in the mathematics cognitive domains*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I.V.S., Martin, M.O., Ruddock, G.J., O'Sullivan, C.Y., Arora, A., & Erberber, E. (2005). *TIMSS 2007 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I.V.S., Martin, M.O., Smith, T.A., Garden, R.A., Gregory, K.D., Gonzalez, E.J., Chrostowki, S.J., & O'Connor, K.M. (2003). *TIMSS assessment frameworks and specifications 2003* (2nd ed.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.