CHAPTER 1

Developing the TIMSS 2015 Achievement Items

Ina V.S. Mullis
Kerry E. Cotter
Bethany G. Fishbein
Victoria A.S. Centurino

Unique Characteristics of TIMSS 2015

The general approach to developing the TIMSS mathematics and science achievement items is similar from assessment cycle to assessment cycle, but each assessment cycle tends to have some unique characteristics that influence the instrument development approach.

- For the first time since 1995, TIMSS and TIMSS Advanced were assessed together in 2015, providing 20 years of trend data for both assessments. TIMSS Advanced is the only international assessment that provides essential information about achievement in advanced mathematics and physics for students in their final year of secondary school. First conducted in 1995 and again in 2008, TIMSS Advanced together with TIMSS 2015 will provide countries with a complete profile of mathematics and science learning from elementary through the end of secondary school.
- TIMSS 2015 was the inaugural year of TIMSS Numeracy. TIMSS Numeracy was introduced in 2015 at the fourth grade to assess fundamental mathematics knowledge, procedures, and problem-solving strategies for students that were likely to find TIMSS 2015 at the fourth grade too difficult.

The TIMSS Approach to Measuring Trends

Because TIMSS is designed to measure trends, the assessments of mathematics and science cannot change dramatically from cycle to cycle. That is, TIMSS is based on a well-known premise for designing trend assessments (ascribed to John Tukey and Albert Beaton):

> "If you want to measure change, do not change the measure."

However, the achievement tests also need to be updated with each cycle to prevent the assessments from becoming dated and no longer relevant to current learning goals. It is important for the content to "keep up with the times" and to be innovative. For example, TIMSS needs

to reflect recent scientific discoveries and to be presented in situations consistent with students' instructional and everyday experiences.

To maintain continuity with past assessments while keeping up with current topics and technology, the TIMSS assessments evolve with each cycle. For assessing mathematics and science, TIMSS has a specific design for the steady release of items after each cycle and replacing them with newly developed items for the following cycle.

Overview of the TIMSS 2015 Achievement Items

Although the majority of the assessment items are carried forward from the previous assessment cycle to measure trends, the task of updating the instruments for each new cycle—every four years for TIMSS since 1995—is a substantial undertaking. Because TIMSS assesses two subjects at two grades, it actually encompasses four different assessments of achievement: mathematics at the fourth and eighth grades and science at the fourth and eighth grades. The two TIMSS 2015 fourth grade assessments required developing and field testing 287 new items, and the two eighth grade assessments required developing and field testing 354 new items. TIMSS Numeracy, the new assessment added at the fourth grade, required developing and field testing 151 items.

The Item Development Process

The TIMSS & PIRLS International Study Center at Boston College uses a collaborative process to develop the new items needed for each TIMSS cycle. A broad overview of the process includes:

- Updating the frameworks for the upcoming assessment
- Developing items and their scoring guides in accordance with the frameworks
- Conducting a full-scale field test
- Selecting the new assessment items based on the frameworks, field test results, and existing items from previous cycles
- Conducting training in how to reliably score responses to constructed response items (i.e., questions to which students provide a written response rather than choosing from a set of options).

The development process is directed and managed by the staff of the TIMSS & PIRLS International Study Center at Boston College, who collectively have considerable experience in the measurement and assessment of mathematics and science achievement. For TIMSS 2015, Executive Director, Ina Mullis, and Assistant Director of Mathematics, Kerry Cotter, managed the mathematics assessment development. Executive Director, Michael Martin, and Associate Director of Science, Victoria Centurino, managed the science assessment development.

Also playing a key role in achievement item development were the National Research Coordinators (NRCs) designated by their countries to be responsible for the complex tasks involved in implementing TIMSS in their countries. The TIMSS & PIRLS International Study Center worked with the NRCs and experts from the countries to develop the new test items including the scoring guides for constructed response items. The NRCs also reviewed the items prior to the field test and helped select the items for the assessment after the field test.

The TIMSS & PIRLS International Study Center prepares an international version of all the TIMSS assessment items in English. Subsequently, the items are translated by participating countries into their languages of instruction with the goal of creating high quality translations that are appropriately adapted for the national context and at the same time are internationally comparable. Therefore, a significant portion of the development and review effort by NRCs is dedicated to ensuring that the test items can be translated accurately.

To provide additional subject-matter expertise and support, external mathematics and science specialists consulted very closely with staff on the development activities. The TIMSS 2015 Chief Mathematics Consultant was Liv Sissel Gronmo, University of Oslo, ILS, Norway, and the TIMSS 2015 Chief Science Consultant was Lee Jones, United States.

Additional advice and guidance were provided through periodic reviews by the Science and Mathematics Review Committee (SMIRC). The SMIRC members for each TIMSS cycle are nominated by countries participating in TIMSS and provide guidance in developing the TIMSS assessments. The TIMSS 2015 SMIRC consisted of 16 members: 6 experts in mathematics and mathematics education and 10 experts in science and science education. It is necessary to have more science members to ensure expertise across the fields of biology, chemistry, and physics. During busy periods, two SMIRC committee members, Mary Lindquist for mathematics and Gerald Wheeler for science, served as advisors to assist in completing specific tasks, such as drafting updated mathematics and science content frameworks and updating scoring guides after the field test.

SMIRC members met four times for TIMSS 2015. At the 1st SMIRC meeting in Oslo, Norway (April 2013), SMIRC reviewed the mathematics and science content frameworks and developed prototype field test items. At the 2nd meeting in St. Petersburg, Russia (September 2013), SMIRC reviewed draft field test items, together with their scoring guides. At the 3rd meeting in Sofia, Bulgaria (July 2014), SMIRC reviewed field test results and made recommendations to the NRCs regarding which items to include in the 2015 mathematics and science assessments. At the final meeting in Seoul, Korea (May 2016), SMIRC conducted the TIMSS 2015 scale anchoring process. Exhibit 1.1 lists the TIMSS 2015 SMIRC members.

Exhibit 1.1: TIMSS 2015 Science and Mathematics Item Review Committee (SMIRC)

Mathematics

Kiril Bankov
Faculty of Mathematics and Informatics
University of Sofia
Bulgaria

Sean Close
Educational Research Centre
St. Patrick's College
Ireland

Khattab Mohammad Ahmad Abulibdeh
National Center for Human Resources Development
Jordan

Sun Sook Noh
College for Education
Ewha Womans University
Korea

Torgeir Onstad
Department of Teacher Education and School
University of Olso, ILS
Norway

Mary Lindquist
United States

Science

Jouni Viiri
Department of Teacher Education
University of Jyväskylä
Finland

Alice Wong
Faculty of Education
University of Hong Kong
Hong Kong SAR

Berenice Michels
National Institute for Curriculum Development
The Netherlands

Newman Burdett
National Foundation for Educational Research
England

Galina Kovaleva
Institute of Content and Methods Education
Russian Academy of Education
Russian Federation

Vitaly Gribov
Physics Faculty
Moscow Lomonosov State University
Russian Federation

Gorazd Planinšič
Faculty of Mathematics and Physics
University of Ljubljana
Slovenia

Wolfgang Dietrich
National Agency for Education
Sweden

Christopher Lazzaro
The College Board
United States

Gerald Wheeler
National Science Teachers' Association
United States

Updating the Mathematics and Science Assessment Frameworks for TIMSS 2015

Updating each TIMSS assessment for 2015 began with reviewing and modifying the assessment frameworks that specify the content to be assessed. The first two chapters of the *TIMSS 2015 Assessment Frameworks*, respectively, describe the mathematics and science frameworks in detail.

The basic structure of the TIMSS mathematics and TIMSS science assessment frameworks is based on two dimensions: content and cognitive. The content domains for mathematics at the fourth grade are number, geometric shapes and measures, and data display. The modified content domains for Numeracy are whole numbers, fractions and decimals, and shapes and measures. At the eighth grade, the mathematics content domains are number, algebra, geometry, and data and chance. For science, the content domains at the fourth grade are life science, physical science, and earth science; at the eighth grade, they are biology, chemistry, physics, and earth science.

Separately for the fourth and eighth grades, the TIMSS mathematics and science frameworks specify several topic areas within each content domain. For example, the algebra content domain contains three topic areas: expressions and operations, equations and inequalities, and relationships and functions. The cognitive domains are the same for mathematics and science: knowing, applying, and reasoning. However, the descriptions of the cognitive skills to be assessed differ somewhat between mathematics and science.

For TIMSS 2015, the mathematics and science frameworks were updated to better reflect the curricula and standards of the countries participating in TIMSS using information from the *TIMSS 2011 Encyclopedia*. These updates were discussed by the NRCs from the participating countries at their first meeting. Following the discussion at the 1st NRC meeting, the NRCs consulted with their national experts and responded to a topic-by-topic survey about how best to update the content and cognitive domains for TIMSS 2015. Next, SMIRC reviewed and revised the frameworks. Using an iterative process, the frameworks as revised by the SMIRC were once again reviewed by the NRCs and updated a final time prior to publication.

Recommendations for updating content and cognitive domains can involve modifying content areas and their weightings (but no more than 5 percent); adding, deleting, or modifying topics within content areas to keep current with research findings and ensure that the number of topics reflects the content area weighting; rewriting to improve clarity for item writers; and perhaps combining some topic areas to reduce redundancy. New for 2015, a new section was added to the science frameworks that describes the science practices to be addressed in science assessments at the fourth and eighth grades. Beyond that, there were no changes in the weighting of content areas for either mathematics or science and only minor revisions to content area topics. The TIMSS 2015 Development schedule is presented in Exhibit 1.2.

Exhibit 1.2: TIMSS 2015 Development Schedule for Achievement Items

Date(s)		Group and Activity
July – December	2012	TIMSS & PIRLS International Study Center conducted content analysis of the curricular topics described in the *TIMSS 2011 Encyclopedia*
October	2012	Task Force proposed updates for the 2015 Assessment Frameworks, incorporating results from the content analysis (Boston, USA)
January	2013	TIMSS & PIRLS International Study Center compiled proposed updates to Assessment Frameworks in preparation for the 1st National Research Coordinator (NRC) meeting
February	2013	NRCs reviewed proposed updates to Assessment Frameworks at 1st NRC meeting (Hamburg, Germany)
March	2013	TIMSS & PIRLS International Study Center incorporated feedback from 1st NRC meeting to further refine the *TIMSS 2015 Assessment Frameworks* and surveyed NRCs online about proposed assessment topic areas and objectives
April	2013	Science and Mathematics Item Review Committee (SMIRC) reviewed proposed mathematics and science frameworks, developed innovative reasoning tasks and prototype items, and reviewed draft *TIMSS 2015 Item Writing Guidelines* at the 1st SMIRC meeting (Oslo, Norway)
May	2013	TIMSS & PIRLS International Study Center prepared final drafts of TIMSS 2015 mathematics and science assessment frameworks, incorporating SMIRC and NRC comments
May	2013	TIMSS & PIRLS International Study Center updated *TIMSS 2015 Item Writing Guidelines*
May	2013	NRCs reviewed *TIMSS 2015 Assessment Frameworks* and developed draft field test items using *TIMSS 2015 Item Writing Guidelines* at 2nd NRC meeting (Amsterdam, The Netherlands)
June – August	2013	TIMSS & PIRLS International Study Center further refined draft field test items and scoring guides and continued to develop additional items to cover frameworks
July	2013	Science and Mathematics Task Forces reviewed and edited draft field test items and scoring guides, developed additional items to cover the frameworks, and classified items into preferred and alternate sets (Boston, USA)
September	2013	SMIRC reviewed draft field test items and scoring guides at 2nd SMIRC meeting (St. Petersburg, Russia)
September	2013	TIMSS & PIRLS International Study Center published *TIMSS 2015 Assessment Frameworks*
September – October	2013	TIMSS & PIRLS International Study Center revised draft field test items and scoring guides to address SMIRC comments
November	2013	NRCs reviewed and approved proposed field test items at 3rd NRC meeting (Budapest, Hungary)
November – December	2013	TIMSS & PIRLS International Study Center assembled field test items into assessment blocks
December	2013	TIMSS & PIRLS International Study Center distributed field test achievement booklets to NRCs
January	2014	TIMSS & PIRLS International Study Center collected student responses to constructed response items from English-speaking countries to develop scoring training materials

Exhibit 1.2: TIMSS 2015 Development Schedule for Achievement Items (Continued)

Date(s)		Group and Activity
February	2014	Science and Mathematics Task Forces modified scoring guides for constructed response items based on student responses and developed scoring training materials for 4th NRC meeting (Boston, USA)
March – April	2014	Countries conducted TIMSS 2015 field test
March	2014	NRCs received scoring training for TIMSS 2015 constructed response field test items at 4th NRC meeting (Sydney, Australia)
April – May	2014	Countries submitted field test achievement data for analysis and review
June	2014	Science and Mathematics Task Forces reviewed field test item statistics
June – July	2014	TIMSS & PIRLS International Study center assembled proposed item blocks in preparation for the 3rd SMIRC meeting
July	2014	SMIRC reviewed proposed item blocks in conjunction with field test results at 3rd SMIRC meeting (Sofia, Bulgaria)
August	2014	NRCs reviewed and approved item blocks for TIMSS 2015 data collection at 5th NRC meeting
August	2014	TIMSS & PIRLS International Study Center distributed TIMSS 2015 data collection achievement booklets to NRCs
October – December	2014	Southern Hemisphere countries conducted TIMSS 2015 data collection
October	2014	TIMSS & PIRLS International Study Center updated and prepared materials for TIMSS 2015 constructed response item scoring training
November	2014	NRCs from Southern Hemisphere countries received scoring training for constructed response items (Wellington, New Zealand)
November	2014	TIMSS & PIRLS International Study Center finalized scoring guides and training materials for constructed response items and distributed them to NRCs
March	2015	NRCs from Northern Hemisphere countries received scoring training for constructed response items at 6th NRC meeting (Prague, Czech Republic)
March – June	2015	Northern Hemisphere countries conducted TIMSS 2015 data collection

Writing and Reviewing the TIMSS 2015 Field Test Items and Scoring Guides

The TIMSS & PIRLS International Study Center uses a collaborative process involving the participating countries to develop test items and scoring guides for the field tests. Most of the 2nd TIMSS NRC meeting in Amsterdam was devoted to a workshop for developing the field test items. The NRCs, together with experienced item writers from participating countries and staff from the TIMSS & PIRLS International Study Center, created the newly developed items for the mathematics and science field tests.

Prior to the workshop, TIMSS & PIRLS International Study Center staff members identified the scope of the item writing task for the field test, examining the weight given to each topic in each of the updated frameworks. Considerations included the total items needed based on the percentage

of weight assigned to a particular area (for example, geometric measurement) in the *TIMSS 2015 Assessment Frameworks*, and the number of topics in that area (two, for example), as well as how many items existed from previous assessments. Because the TIMSS & PIRLS International Study Center generally field tests twice the number of items actually required, the field test included the target number of new items needed multiplied by two. For TIMSS 2015, about 800 items were field tested (see Exhibit 1.4).

In preparation for the item writing workshop, the TIMSS & PIRLS International Study Center updated the *TIMSS 2015 Item Writing Guidelines*, an item writing manual specifically developed for TIMSS assessments. The *Item Writing Guidelines* contain general information about procedures for obtaining good measurement (for instance, items should be independent and not provide clues to the correct responses of other items) as well as specific information on how to deal with translation and comparability issues (for example, using TIMSS' fictitious unit of currency, the "zed," for any money items). The *Item Writing Guidelines* include the necessary steps for developing scoring guides, as well as checklists for reviewing TIMSS items.

At the TIMSS item writing workshop, country representatives were divided into teams and given specific item writing assignments to ensure that enough field test items were developed in each of the content areas and cognitive processes areas specified in the frameworks. The TIMSS & PIRLS International Study Center staff and consultants used the *Item Writing Guidelines* to provide training to the teams on item writing procedures for the TIMSS assessments. Once teams had completed their item writing assignments, each team reviewed the items drafted by other teams. In addition, some teams continued to send items to the TIMSS & PIRLS International Study Center for several weeks after the item writing workshop. Exhibit 1.3 shows the number of participants in the TIMSS 2015 item writing workshop and the number of items written.

Exhibit 1.3: TIMSS 2015 Item Writing Workshop to Develop Field Test Items

Attendees	
Number of Countries and Benchmarking Entities	45
Number of Country Representatives	114
Approximate Number of Field Test Items Written at Item Writing Workshop	
Fourth Grade Mathematics	160
Eighth Grade Mathematics	200
Fourth Grade Science	160
Eighth Grade Science	270

Following the item writing workshop, the draft set of field test items received a thorough review by the TIMSS & PIRLS International Study Center. Reviewers included staff, the chief consultants, and consultants experienced in developing assessment items, such as those from Educational Testing Service, the National Foundation for Educational Research in England, and the Australian Council for Educational Research, as well as SMIRC members with particular item writing skills.

Finally, the proposed field test blocks were reviewed by the TIMSS 2015 SMIRC and NRCs prior to field test instrument production. The TIMSS & PIRLS International Study Center implemented the suggested revisions and provided the final international version of the field test booklets to the NRCs so that they could begin translating the field test materials into their languages of instruction.

The TIMSS 2015 Field Test

The TIMSS field test followed typical TIMSS procedures, where it served as a full-scale "dress rehearsal" operationally for the assessment. That is, the data collection and scoring procedures to be employed in the assessment were practiced in the field test. In addition, the field test provided important information about how well each prospective item functioned and provided a basis for selecting items for the assessment.

The field test was designed to be conducted for approximately 30 schools in each country and yield at least 200 student responses to each mathematics and science item. Generally, the samples for the field test and the assessment are drawn simultaneously, using the same random sampling procedures. This ensures that field test samples closely approximate assessment samples, and that a school is selected for either the field test or the assessment, but not both. For example, if 150 schools are needed for the assessment and another 30 for the field test, then a larger sample of 180 schools is selected and a systematic sample of 30 schools is selected from the 180 schools.

The TIMSS 2015 field test was conducted in March–April 2014. Exhibits 1.4 through 1.8 provide a detailed summary of the field test effort, including the number of students, teachers, and schools that participated, and the number of items listed by format, content domain, and cognitive domain. Approximately 10,000 student responses from more than 40 countries per grade were used to evaluate the measurement properties of each field test assessment item.

Exhibit 1.4: Overview of the TIMSS 2015 Field Test

	Fourth Grade	Numeracy	Eighth Grade
Items			
Mathematics	147	151	182
Science	140		172
Total	**287**	**151**	**354**
Responses per Item (approx.)	200	200	200
Participants			
Countries	43	7	39
Benchmarking Entities	5	0	4
Students	54,679	4,522	51,994
Teachers	3,772	296	6,097
Schools	1,469	164	1,142

Exhibit 1.5: TIMSS 2015 Number of Field Test Items by Content Domain and Item Format – Fourth Grade

Content Domain	Number of Multiple-Choice Items	Number of Constructed Response Items	Total Number of Items	Total Number of Score Points	Percentage of Score Points
Mathematics Items					
Number	42	46	88	93	60%
Geometric Shapes and Measures	23	18	41	43	28%
Data Display	3	15	18	18	12%
Total	**68**	**79**	**147**	**154**	
Mathematics - Numeracy Items					
Whole Numbers	33	42	75	77	50%
Fractions and Decimals	14	11	25	25	16%
Shapes and Measures	26	25	51	52	34%
Total	**73**	**78**	**151**	**154**	
Science Items					
Life Science	28	34	62	66	45%
Physical Science	31	20	51	53	36%
Earth Science	21	6	27	28	19%
Total	**80**	**60**	**140**	**147**	

Exhibit 1.6: TIMSS 2015 Number of Field Test Items by Cognitive Domain and Item Format – Fourth Grade

Cognitive Domain	Number of Multiple-Choice Items	Number of Constructed Response Items	Total Number of Items	Total Number of Score Points	Percentage of Score Points
Mathematics Items					
Knowing	33	25	58	58	38%
Applying	24	35	59	63	41%
Reasoning	11	19	30	33	21%
Total	**68**	**79**	**147**	**154**	
Mathematics - Numeracy Items					
Knowing	39	25	64	64	42%
Applying	25	35	60	61	40%
Reasoning	9	18	27	29	19%
Total	**73**	**78**	**151**	**154**	
Science Items					
Knowing	33	20	53	56	38%
Applying	29	28	57	59	40%
Reasoning	18	12	30	32	22%
Total	**80**	**60**	**140**	**147**	

Exhibit 1.7: TIMSS 2015 Number of Field Test Items by Content Domain and Item Format – Eighth Grade

Content Domain	Number of Multiple-Choice Items	Number of Constructed Response Items	Total Number of Items	Total Number of Score Points	Percentage of Score Points
Mathematics Items					
Number	19	28	47	51	26%
Algebra	24	24	48	51	26%
Geometry	21	25	46	51	26%
Data and Chance	20	21	41	45	23%
Total	**84**	**98**	**182**	**198**	
Science Items					
Biology	31	29	60	72	37%
Chemistry	15	21	36	38	20%
Physics	24	19	43	46	24%
Earth Science	20	13	33	36	19%
Total	**90**	**82**	**172**	**192**	

Exhibit 1.8: TIMSS 2015 Number of Field Test Items by Cognitive Domain and Item Format – Eighth Grade

Cognitive Domain	Number of Multiple-Choice Items	Number of Constructed Response Items	Total Number of Items	Total Number of Score Points	Percentage of Score Points
Mathematics Items					
Knowing	35	12	47	47	24%
Applying	32	46	78	83	42%
Reasoning	17	40	57	68	34%
Total	**84**	**98**	**182**	**198**	
Science Items					
Knowing	46	18	64	75	39%
Applying	32	37	69	74	39%
Reasoning	12	27	39	43	22%
Total	**90**	**82**	**172**	**192**	

Developing the Materials for TIMSS 2015 Field Test Scoring Training

It is necessary to prepare scoring training materials for the newly developed constructed response field test items in advance of the field test so field test scoring can occur immediately upon completion of data collection. To provide "grist" for these materials, Canada, Ireland, and Singapore, administered the newly developed constructed response field test items in a small selection of classrooms with English-speaking students. Pilot materials were completed in December 2013 and responses were gathered from students in January 2014. The goal was to collect a total of approximately 200 responses to each newly developed constructed response field test item to provide example student responses in the field test scoring guides and sets of training materials. Exhibit 1.9 provides the number of items included in the pilot test and the number of student responses collected.

Exhibit 1.9: Pilot Test Student Responses for Field Test Scoring Training Materials Development

	Fourth Grade	Eighth Grade
Items		
Mathematics	24	24
Science	49	76
Total	**73**	**100**
Responses per Item (approx.)	180	160
Participants		
Countries	Canada, Ireland, Singapore	Canada, Ireland, Singapore
Number of Students (approx.)	360	320

Additionally, the United States arranged for cognitive labs in Washington, D.C. and California. Each TIMSS constructed response item was presented to approximately five students, who were observed and prompted to answer questions about the clarity, difficulty, and familiarity of the item content and format. The TIMSS & PIRLS International Study Center received the cognitive lab reports in February 2014. Exhibit 1.10 provides the number of items included in the cognitive labs and the number of student responses collected.

Exhibit 1.10: Cognitive Lab Student Responses

	Fourth Grade	Eighth Grade
Mathematics Items	20	20
Science Items	20	20
Total Items	**40**	**40**
Responses per Item (approx.)	5	5
Number of Students (approx.)	50	50

The TIMSS 2015 NRCs and their scoring supervisors received scoring training for the field test constructed response items in March 2014 in Sydney, Australia, as part of the 4th TIMSS 2015 NRC Meeting. Sets of example and practice papers were created for 34 fourth grade items and 33 eighth grade items. The example and practice paper sets for each item included a scoring guide, approximately 8–10 example papers illustrating the categories in the scoring guide, and approximately 8–10 practice papers so that country representatives could practice making distinctions among categories and reach agreement about how to make consistent scoring decisions across countries.

At the scoring training sessions, the trainers explained the purpose of each item and read it aloud. The trainer then described the scoring guide, explaining each category and the rationale

for the score given to each example paper. After the country representatives scored the practice papers, any inconsistencies in scoring were discussed, and, as necessary, the field test guides were clarified and sometimes categories were revised.

Finalizing the TIMSS 2015 Achievement Items

Subsequent to the field test, the TIMSS & PIRLS International Study Center analyzed the TIMSS field test data and prepared almanacs containing summary item statistics for each field test item. The data almanac for an item contained, row by row for each country: the sample size, the item difficulty and discrimination, the percentage of students answering each option (multiple-choice) or in each score category (constructed response), the point-biserial correlation for each multiple-choice option or constructed response category, and the degree of scoring agreement for constructed response items.

The field test data were used by the TIMSS & PIRLS International Study Center, expert committees, and NRCs to assess the quality of the field test items. The TIMSS & PIRLS International Study Center staff members, together with external consultants, first reviewed the field test data to make an initial judgment about the quality of each item based on its measurement properties (item statistics). Items were eliminated from further consideration if they had poor measurement properties, such as being too difficult or easy or having low discrimination. Particular attention was paid to unusual item statistics in individual countries since these could indicate errors in translation.

After the item-by-item review, the TIMSS & PIRLS International Study Center staff collaborated with consultants to assemble a set of recommended assessment blocks for review by the expert committee (SMIRC). SMIRC members scrutinized the recommendations for the newly developed assessment blocks, reviewing the items and scoring guides for content accuracy, clarity, and adherence to the frameworks. In addition, the newly developed items were considered in relation to the trend item blocks for overall coherence as a complete assessment.

The SMIRC's recommendations were implemented by staff, and the penultimate assessment blocks were sent to the NRCs for review. NRCs had the opportunity to review the recommended materials in light of the field test results and within the security of their own countries. Each country also could check any unusual national results that might be an indication of translation errors and correct the translation as necessary or recommend revisions to accommodate translation. Finally, the 5th NRC meeting held in Paris, France in August 2014 was devoted to reviewing all the newly developed items.

Distribution of TIMSS 2015 Items by Content and Cognitive Domains

Exhibits 1.11 through 1.14 present the number of trend and newly developed items as well as the number of score points in the TIMSS 2015 mathematics and science assessments. The number of items represents the number of distinct questions in the assessment, while the number of score points represents the complexity and weight given to each item.

Exhibit 1.11: TIMSS 2015 Achievement Items by Content Domain – Fourth Grade

Content Domain	Number of Trend Items in TIMSS 2015	Percentage of Trend Score Points	Number of New Items in TIMSS 2015	Percentage of New Score Points	Total Items	Achieved Percentage of Score Points	Target Percentage of Score Points
Mathematics Items							
Number	48 (49)	46%	41 (46)	61%	89 (95)	52%	50%
Geometric Shapes and Measures	37 (38)	36%	19 (21)	28%	56 (59)	32%	35%
Data Display	17 (19)	18%	7 (9)	12%	24 (28)	15%	15%
Total	**102 (106)**		**67 (76)**		**169 (182)**		
Mathematics - Numeracy Items							
Whole Numbers			52 (53)	50%	52 (53)	50%	50%
Fractions and Decimals			15 (15)	14%	15 (15)	14%	15%
Shapes and Measures			35 (38)	36%	35 (38)	36%	35%
Total			**102 (106)**		**102 (106)**		
Science Items							
Life Science	47 (52)	48%	32 (35)	44%	79 (87)	46%	45%
Physical Science	35 (35)	32%	29 (30)	38%	64 (65)	35%	35%
Earth Science	19 (22)	20%	14 (14)	18%	33 (36)	19%	20%
Total	**101 (109)**		**75 (79)**		**176 (188)**		

Score points are shown in parentheses.

Because percentages are rounded to the nearest whole number, some totals may appear inconsistent.

Counts of TIMSS Numeracy achievement items do not include the two fourth grade TIMSS 2015 mathematics blocks (see Chapter 4 of the *TIMSS 2015 Assessment Frameworks*).

Exhibit 1.12: TIMSS 2015 Achievement Items by Cognitive Domain – Fourth Grade

Cognitive Domain	Number of Trend Items in TIMSS 2015	Percentage of Trend Score Points	Number of New Items in TIMSS 2015	Percentage of New Score Points	Total Items	Achieved Percentage of Score Points	Target Percentage of Score Points
Mathematics Items							
Knowing	41 (41)	39%	23 (24)	32%	64 (65)	36%	40%
Applying	42 (45)	42%	30 (35)	46%	72 (80)	44%	40%
Reasoning	19 (20)	19%	14 (17)	22%	33 (37)	20%	20%
Total	**102 (106)**		**67 (76)**		**169 (182)**		
Mathematics - Numeracy Items							
Knowing			55 (55)	52%	55 (55)	52%	50%
Applying			35 (36)	34%	35 (36)	34%	15%
Reasoning			12 (15)	14%	12 (15)	14%	35%
Total			**102 (106)**		**102 (106)**		
Science Items							
Knowing	41 (44)	40%	31 (34)	43%	72 (78)	41%	40%
Applying	40 (43)	39%	27 (28)	35%	67 (71)	38%	40%
Reasoning	20 (22)	20%	17 (17)	22%	37 (39)	21%	20%
Total	**101 (109)**		**75 (79)**		**176 (188)**		

Score points are shown in parentheses.

Because percentages are rounded to the nearest whole number, some totals may appear inconsistent.

Counts of TIMSS Numeracy achievement items do not include the two fourth grade TIMSS 2015 mathematics blocks (see Chapter 4 of the *TIMSS 2015 Assessment Frameworks*).

Exhibit 1.13: TIMSS 2015 Achievement Items by Content Domain – Eighth Grade

Content Domain	Number of Trend Items in TIMSS 2015	Percentage of Trend Score Points	Number of New Items in TIMSS 2015	Percentage of New Score Points	Total Items	Achieved Percentage of Score Points	Target Percentage of Score Points
Mathematics Items							
Number	40 (45)	34%	24 (25)	26%	64 (70)	31%	30%
Algebra	40 (42)	31%	22 (23)	24%	62 (65)	28%	30%
Geometry	22 (22)	16%	21 (25)	26%	43 (47)	21%	20%
Data and Chance	25 (25)	19%	18 (22)	23%	43 (47)	21%	20%
Total	**127 (134)**		**85 (95)**		**212 (229)**		
Science Items							
Biology	47 (51)	38%	28 (36)	34%	75 (87)	36%	35%
Chemistry	26 (27)	20%	18 (19)	18%	44 (46)	19%	20%
Physics	32 (32)	24%	24 (25)	24%	56 (57)	24%	25%
Earth Science	23 (24)	18%	22 (25)	24%	45 (49)	21%	20%
Total	**128 (134)**		**92 (105)**		**220 (239)**		

Score points are shown in parentheses.

Because percentages are rounded to the nearest whole number, some totals may appear inconsistent.

Exhibit 1.14: TIMSS 2015 Achievement Items by Cognitive Domain – Eighth Grade

Cognitive Domain	Number of Trend Items in TIMSS 2015	Percentage of Trend Score Points	Number of New Items in TIMSS 2015	Percentage of New Score Points	Total Items	Achieved Percentage of Score Points	Target Percentage of Score Points
			Mathematics Items				
Knowing	45 (46)	34%	24 (24)	25%	69 (70)	31%	35%
Applying	54 (58)	43%	41 (45)	47%	95 (103)	45%	40%
Reasoning	28 (30)	22%	20 (26)	27%	48 (56)	24%	25%
Total	**127 (134)**		**85 (95)**		**212 (229)**		
			Science Items				
Knowing	40 (41)	31%	37 (44)	42%	77 (85)	36%	35%
Applying	58 (61)	46%	33 (37)	35%	91 (98)	41%	35%
Reasoning	30 (32)	24%	22 (24)	23%	52 (56)	23%	30%
Total	**128 (134)**		**92 (105)**		**220 (239)**		

Score points are shown in parentheses.
Because percentages are rounded to the nearest whole number, some totals may appear inconsistent.

Distribution of TIMSS 2015 Item Formats within Content and Cognitive Domains

Exhibits 1.15 through 1.18 display the number of items (and score points) by item format for each content and cognitive domain. As described in the *TIMSS 2015 Assessment Frameworks*, at least half of the total number of score points represented by all the questions should come from multiple-choice items. Most TIMSS multiple-choice items are worth one score point, although some compound multiple-choice items are worth two score points. The 2-point compound multiple-choice items are scored as all parts answered correctly as fully correct (2 score points), and most parts answered correctly as partially correct (1 score point). Constructed response items generally are worth one or two score points depending on the degree of complexity involved. The 1-point constructed response items are scored as correct (1 score point) or incorrect (0 score points), whereas 2-point constructed response items are scored as fully correct (2 score points), partially correct (1 score point), or incorrect (0 score points). Fully correct responses show a complete or deeper understanding of a task while partially correct responses demonstrate only a partial understanding of the concepts or procedures embodied in the task.

Exhibit 1.15: TIMSS 2015 Achievement Items by Content Domain and Item Format – Fourth Grade

Content Domain	Multiple-Choice Items		Constructed Response Items		Total Items	Percentage of Score Points
	Four Response Options	Compound	1 Point	2 Points		
Mathematics Items						
Number	44 (44)	2 (2)	37 (37)	6 (12)	89 (95)	52%
Geometric Shapes and Measures	35 (35)		18 (18)	3 (6)	56 (59)	32%
Data Display	8 (8)		12 (12)	4 (8)	24 (28)	15%
Total	**87 (87)**	**2 (2)**	**67 (67)**	**13 (26)**	**169 (182)**	
Achieved Percentage of Score Points	**49%**		**51%**			
Target Percentage of Score Points	**50%**		**50%**			
Mathematics – Numeracy Items						
Whole Numbers	21 (21)		30 (30)	1 (2)	52 (53)	50%
Fractions and Decimals	7 (7)		8 (8)		15 (15)	14%
Shapes and Measures	17 (17)	1 (1)	14 (14)	3 (6)	35 (38)	36%
Total	**45 (45)**	**1 (1)**	**52 (52)**	**4 (8)**	**102 (106)**	
Achieved Percentage of Score Points	**43%**		**57%**			
Target Percentage of Score Points	**50%**		**50%**			
Science Items						
Life Science	37 (37)	2 (2)	32 (32)	8 (16)	79 (87)	46%
Physical Science	32 (32)	4 (4)	27 (27)	1 (2)	64 (65)	35%
Earth Science	21 (21)	2 (2)	7 (7)	3 (6)	33 (36)	19%
Total	**90 (90)**	**8 (8)**	**66 (66)**	**12 (24)**	**176 (188)**	
Achieved Percentage of Score Points	**52%**		**48%**			
Target Percentage of Score Points	**50%**		**50%**			

Score points are shown in parentheses.

Because percentages are rounded to the nearest whole number, some totals may appear inconsistent.

Counts of TIMSS Numeracy achievement items do not include the two fourth grade TIMSS 2015 mathematics blocks (see Chapter 4 of the *TIMSS 2015 Assessment Frameworks*).

Exhibit 1.16: TIMSS 2015 Achievement Items by Cognitive Domain and Item Format – Fourth Grade

Cognitive Domain	Multiple-Choice Items		Constructed Response Items		Total Items	Percentage of Score Points
	Four Response Options	Compound	1 Point	2 Points		
Mathematics Items						
Knowing	35 (35)	2 (2)	26 (26)	1 (2)	64 (65)	36%
Applying	36 (36)		28 (28)	8 (16)	72 (80)	44%
Reasoning	16 (16)		13 (13)	4 (8)	33 (37)	20%
Total	**87 (87)**	**2 (2)**	**67 (67)**	**13 (26)**	**169 (182)**	
Achieved Percentage of Score Points	**49%**		**51%**			
Target Percentage of Score Points	**50%**		**50%**			
Mathematics – Numeracy Items						
Knowing	29 (29)	1 (1)	25 (25)		55 (55)	52%
Applying	11 (11)		23 (23)	1 (2)	35 (36)	34%
Reasoning	5 (5)		4 (4)	3 (6)	12 (15)	14%
Total	**45 (45)**	**1 (1)**	**52 (52)**	**4 (8)**	**102 (106)**	
Achieved Percentage of Score Points	**43%**		**57%**			
Target Percentage of Score Points	**50%**		**50%**			
Science Items						
Knowing	42 (42)	5 (5)	19 (19)	6 (12)	72 (78)	41%
Applying	31 (31)	1 (1)	31 (31)	4 (8)	67 (71)	38%
Reasoning	17 (17)	2 (2)	16 (16)	2 (4)	37 (39)	21%
Total	**90 (90)**	**8 (8)**	**66 (66)**	**12 (24)**	**176 (188)**	
Achieved Percentage of Score Points	**52%**		**48%**			
Target Percentage of Score Points	**50%**		**50%**			

Score points are shown in parentheses.

Because percentages are rounded to the nearest whole number, some totals may appear inconsistent.

Counts of TIMSS Numeracy achievement items do not include the two fourth grade TIMSS 2015 mathematics blocks (see Chapter 4 of the *TIMSS 2015 Assessment Frameworks*).

Exhibit 1.17: TIMSS 2015 Achievement Items by Content Domain and Item Format – Eighth Grade

Content Domain	Multiple-Choice Items		Constructed Response Items		Total Items	Percentage of Score Points
	Four Response Options	Compound	1 Point	2 Points		
Mathematics Items						
Number	28 (28)	1 (1)	29 (29)	6 (12)	64 (70)	31%
Algebra	35 (35)		24 (24)	3 (6)	62 (65)	28%
Geometry	22 (22)		17 (17)	4 (8)	43 (47)	21%
Data and Chance	27 (27)	2 (4)	12 (12)	2 (4)	43 (47)	21%
Total	**112 (112)**	**3 (5)**	**82 (82)**	**15 (30)**	**212 (229)**	
Achieved Percentage of Score Points	**51%**		**49%**			
Target Percentage of Score Points	**50%**		**50%**			
Science Items						
Biology	35 (35)	1 (1)	27 (27)	12 (24)	75 (87)	36%
Chemistry	19 (19)	4 (5)	20 (20)	1 (2)	44 (46)	19%
Physics	31 (31)	2 (3)	23 (23)		56 (57)	24%
Earth Science	26 (26)	3 (4)	13 (13)	3 (6)	45 (49)	21%
Total	**111 (111)**	**10 (13)**	**83 (83)**	**16 (32)**	**220 (239)**	
Achieved Percentage of Score Points	**52%**		**48%**			
Target Percentage of Score Points	**50%**		**50%**			

Score points are shown in parentheses.
Because percentages are rounded to the nearest whole number, some totals may appear inconsistent

Exhibit 1.18: TIMSS 2015 Achievement Items by Cognitive Domain and Item Format – Eighth Grade

Cognitive Domain	Multiple-Choice Items		Constructed Response Items		Total Items	Percentage of Score Points
	Four Response Options	Compound	1 Point	2 Points		
Mathematics Items						
Knowing	49 (49)	1 (1)	18 (18)	1 (2)	69 (70)	31%
Applying	48 (48)		39 (39)	8 (16)	95 (103)	45%
Reasoning	15 (15)	2 (4)	25 (25)	6 (12)	48 (56)	24%
Total	**112 (112)**	**3 (5)**	**82 (82)**	**15 (30)**	**212 (229)**	
Achieved Percentage of Score Points	**51%**		**49%**			
Target Percentage of Score Points	**50%**		**50%**			
Science Items						
Knowing	59 (59)	5 (7)	7 (7)	6 (12)	77 (85)	36%
Applying	39 (39)	5 (6)	41 (41)	6 (12)	91 (98)	41%
Reasoning	13 (13)		35 (35)	4 (8)	52 (56)	23%
Total	**111 (111)**	**10 (13)**	**83 (83)**	**16 (32)**	**220 (239)**	
Achieved Percentage of Score Points	**52%**		**48%**			
Target Percentage of Score Points	**50%**		**50%**			

Score points are shown in parentheses.
Because percentages are rounded to the nearest whole number, some totals may appear inconsistent.

TIMSS 2015 Constructed Response Scoring Training

In preparation for the main data collection scoring training, some TIMSS 2015 scoring guides were further refined or clarified based on the results of the field test. This also included a thorough review of the field test scoring training materials to ensure that the student responses were still suitable for the updated scoring guides. In some cases, example and practice sets used in the field test were expanded to further illustrate particular aspects of a scoring guide. For TIMSS 2015 scoring training, the example and practice paper training sets included those used in TIMSS 2011 for the trend items and the updated training sets for the newly developed items selected for TIMSS 2015, resulting in 27 example and practice paper sets for fourth grade and 29 for eighth grade.

To provide scoring training for all the countries participating in TIMSS 2015, the TIMSS & PIRLS International Study Center conducted two training sessions. First, the NRCs for Southern Hemisphere countries and their scoring supervisors received scoring training in November 2014 in Wellington, New Zealand. NRCs for Northern Hemisphere countries and their scoring supervisors received scoring training in March 2015 in Prague, Czech Republic as part of the 6th TIMSS 2015 NRC Meeting. Exhibit 1.19 shows the number of participants in the two scoring training sessions.

Exhibit 1.19: TIMSS 2015 Scoring Training Participation

Participants	Southern Hemisphere	Northern Hemisphere
Number of Countries	8	58
Number of Benchmarking Entities	1	5
Number of Country Representatives	32	152

The Process Following Instrument Development

In general, after the participating countries received the international version of the assessment instruments, they began the process of translation and cultural adaptation (some adaptation to local usage typically is necessary even in English-speaking countries) and production of the materials for printing. At the same time, countries made final arrangements for data collection, including the host of activities necessary to obtain school participation, implement test administration, and score the responses to the tests and questionnaires (see following chapters).