CHAPTER 4

TIMSS Advanced 2015 Assessment Design

Michael O. Martin, Ina V.S. Mullis, and Pierre Foy

Overview

The TIMSS Advanced 2015 assessment measures trends in student achievement in advanced mathematics and physics at the end of secondary schooling for students with advanced preparation in these subjects. The assessment comprises written tests in advanced mathematics and physics together with sets of questionnaires that gather information on the educational and social contexts for achievement at the end of secondary schooling. First administered in 1995 and again in 2008, TIMSS Advanced 2015 continues this trend line for those countries that participated in prior assessments, with each assessment linked to the next. Significantly in 2015, and for the first time since 1995, TIMSS Advanced will be administered in the same year as the fourth and eighth grade TIMSS assessments of mathematics and science; this will enable countries participating at all three levels (the fourth grade, the eighth grade, and at the end of secondary school) to collect data on student achievement in mathematics and science spanning the entire primary and secondary education system.

As described in the advanced mathematics and physics assessment frameworks (Chapters 1 and 2, respectively), the TIMSS Advanced assessments are wide ranging in their coverage of these subjects; the assessments are designed to provide valid and reliable information on the full range of student proficiency in each subject, as well as in the major content and cognitive domains.

A consequence of the TIMSS Advanced reporting goals is that the assessments require more assessment items than can reasonably be given to a student in the available testing time. Accordingly, TIMSS Advanced 2015 uses a matrix-sampling approach that assembles the pool of achievement

items in advanced mathematics and physics into a set of 12 assessment booklets—6 advanced mathematics, and 6 physics booklets—with each student completing one booklet only. Each item appears in two booklets, providing a mechanism for linking the student responses from the various booklets. Booklets are distributed among sampled students so that the groups of students responding to each booklet are approximately equivalent in terms of student ability.

Student Populations Assessed

TIMSS Advanced assesses the advanced mathematics and physics achievement of students in the final year of secondary schooling. This is the twelfth year of formal schooling in most countries. The target populations for the TIMSS Advanced assessments are defined as follows:

> For advanced mathematics, all students in the final year of secondary schooling who are taking advanced mathematics courses.
>
> For physics, all students in the final year of secondary schooling who are taking physics courses.

Student eligibility is determined in terms of the courses the student has taken and, in countries with tracked educational systems, the track to which the student belongs. The decision as to which mathematics or physics courses should be included in defining the target population is determined by each participating country. In general, the courses included should be those taken by the most advanced students, typically those students planning further study in mathematics or physics at university or other institutes of higher education. Courses should cover most of the advanced mathematics and physics content topics specified in Chapters 1 and 2 of the *TIMSS Advanced 2015 Assessment Frameworks*. Depending on their course experience, students in the final year of secondary schooling may belong to the advanced mathematics target population, the physics target population, or both. Students who belong to both populations will be randomly assigned either an advanced mathematics booklet or a physics booklet.

Reporting Student Achievement

TIMSS Advanced 2015 will provide a detailed picture of the achievement of advanced mathematics and physics students in the final year of secondary school in each participating country. This will include achievement in each of the content and cognitive domains (as defined in Chapters 1 and 2) as well as overall advanced mathematics and physics achievement. Consistent with the goal of providing valid and reliable information on the full range of student proficiency in each subject, the complete TIMSS Advanced 2015 assessment consists of a large pool of advanced mathematics and physics questions, known as items. However, in order to keep the assessment burden on any one student to a minimum, each student is presented with only a sample of the items, as described in the next section. Following data collection, student responses are placed on common advanced mathematics and physics scales in order to provide an overall picture of the assessment results for each country.

One of the strengths of TIMSS Advanced is its measurement of trends over time in advanced mathematics and physics achievement. The TIMSS Advanced achievement scales provide a common metric on which countries can compare the progress of their student populations in advanced mathematics and physics from assessment to assessment. The TIMSS Advanced achievement scales were established in 1995, separately for advanced mathematics and physics, so that 100 points on the scale was equal to one standard deviation across all of the countries that participated in TIMSS 1995, and the scale midpoint of 500 was equal to the international average across those countries. Using items that were administered in both 1995 and 2008 assessments as a basis for linking the two sets of assessment results, the TIMSS Advanced 2008 data also were placed on the scale so that countries could gauge changes in students' advanced mathematics and physics achievement since 1995. Using similar procedures, the data from TIMSS Advanced 2015 will be placed on the TIMSS Advanced scales, enabling TIMSS Advanced 2015 countries that have participated in previous assessments to have comparable achievement data from 1995, 2008, and 2015, and to plot changes in performance over this 20-year period.

As previously mentioned, in addition to the achievement scales for advanced mathematics and physics overall, TIMSS Advanced 2015 includes scales for reporting relative student performance in each of the advanced mathematics and physics content and cognitive domains. Specifically, in advanced mathematics there are three content scales, corresponding to three

content domains: algebra, calculus, and geometry. Similarly, in physics there are also three content scales, corresponding to three content domains: mechanics and thermodynamics, electricity and magnetism, and wave phenomena and atomic/nuclear physics. The *TIMSS Advanced 2015 Assessment Frameworks* specify three cognitive domains—knowing, applying, and reasoning—which span the content of both advanced mathematics and physics, and for which reporting scales are constructed.

TIMSS Advanced 2015 Student Booklet Design

A consequence of the ambitious reporting goals of TIMSS Advanced is that many more items are required for the assessment than can be answered by any one student in the available testing time. In order to address this challenge, TIMSS Advanced 2015 uses a matrix-sampling approach: the entire assessment pool of advanced mathematics and physics items are packaged into a set of 6 advanced mathematics booklets and 6 physics booklets, with each student completing just one booklet. Each item appears in two booklets, providing a mechanism for linking together the student responses from the various booklets. Booklets are distributed randomly among students in participating classrooms so that the groups of students completing each booklet are approximately equivalent in terms of ability. TIMSS Advanced uses item response theory scaling methods in order to assemble a comprehensive picture of the achievement of a country's student population by pooling individual students' responses to the booklets that they are assigned. This approach reduces to manageable proportions what otherwise would be an impossible student burden, albeit at the cost of greater complexity in booklet assembly, data collection, and data analysis.

In order to facilitate the process of creating the student achievement booklets, TIMSS Advanced groups the assessment items into a series of item blocks, with each item block consisting of approximately 10 items and requiring 30 minutes of assessment time. As far as possible, within each block the distribution of items across content and cognitive domains matches the distribution across the overall item pool. TIMSS Advanced 2015 consists of 18 item blocks in total: 9 blocks of advanced mathematics items, and 9 blocks of physics items. This represents an increase of 4 blocks over the 14 blocks that formed the basis of TIMSS Advanced 2008. The additional item blocks were added in order to provide more extensive coverage of the content and cognitive domains. Student booklets for advanced mathematics and physics are assembled from various combinations of these item blocks.

Following the 2008 assessment, three of the advanced mathematics item blocks and three of the physics blocks were retained and kept secure for use in measuring trends in 2015. The remaining 8 blocks (4 advanced mathematics, and 4 physics) were released into the public domain for use in publications, research, and teaching, to be replaced by newly-developed items for the TIMSS Advanced 2015 assessment. Accordingly, the 18 blocks in the TIMSS Advanced 2015 assessment comprise 6 blocks of trend items (3 advanced mathematics, and 3 physics) and 12 blocks of items newly developed for 2015. As presented in Exhibit 3, the TIMSS Advanced 2015 advanced mathematics blocks are labeled M1 through M9, and the physics blocks P1 through P9.

Exhibit 3: TIMSS Advanced 2015 Item Block Design

Advanced Mathematics Blocks		Physics Blocks	
M1	Block M2 from TIMSS Advanced 2008	P1	Block P2 from TIMSS Advanced 2008
M2	New items for TIMSS Advanced 2015	P2	New items for TIMSS Advanced 2015
M3	Block M4 from TIMSS Advanced 2008	P3	Block P4 from TIMSS Advanced 2008
M4	New items for TIMSS Advanced 2015	P4	New items for TIMSS Advanced 2015
M5	Block M5 from TIMSS Advanced 2008	P5	Block P5 from TIMSS Advanced 2008
M6	New items for TIMSS Advanced 2015	P6	New items for TIMSS Advanced 2015
M7	New items for TIMSS Advanced 2015	P7	New items for TIMSS Advanced 2015
M8	New items for TIMSS Advanced 2015	P8	New items for TIMSS Advanced 2015
M9	New items for TIMSS Advanced 2015	P9	New items for TIMSS Advanced 2015

Students are expected to spend, on average, 30 minutes on each item block. Consequently, the 9 blocks of advanced mathematics items are estimated to contain 4½ hours of testing time, and the physics blocks a further 4½ hours. From past experience with TIMSS Advanced, National Research Coordinators from participating countries agreed that the testing time for any one student should not be increased from previous assessments; thus, as in the past, the assessment time for each student booklet (advanced mathematics or physics) must fit into 90 minutes. An additional 30 minutes for a student questionnaire also is required.

In choosing how to distribute item blocks across student achievement booklets, the major goal was to maximize coverage of the framework while ensuring that every student responded to sufficient items in order to provide reliable measurement of trends in both advanced mathematics and physics.

A further goal was to ensure that achievement in the advanced mathematics and physics content and cognitive domains could be measured reliably. In order to enable linking among booklets while keeping the number of booklets to a minimum, each item block appears in two booklets.

As presented in Exhibit 4, the 18 assessment item blocks are distributed across 12 student achievement booklets. Booklets 1 through 6 contain advanced mathematics items, and Booklets 7 through 12 contain physics items. Each student booklet consists of three item blocks.

Exhibit 4: TIMSS Advanced 2015 Student Achievement Booklet Design

	Assessment Blocks		
Student Achievement Booklet	**Advanced Mathematics**		
	Part 1	**Part 2**	**Part 3**
Booklet 1	M1	M2	M4
Booklet 2	M4	M3	M6
Booklet 3	M6	M7	M5
Booklet 4	M3	M8	M7
Booklet 5	M8	M5	M9
Booklet 6	M2	M9	M1
	Physics		
Booklet 7	P1	P2	P4
Booklet 8	P4	P3	P6
Booklet 9	P6	P7	P5
Booklet 10	P3	P8	P7
Booklet 11	P8	P5	P9
Booklet 12	P2	P9	P1

Countries participating in TIMSS Advanced aim for a sample of approximately 3,600 advanced mathematics students and the same number of physics students in order to ensure that there are enough respondents for each item. In classes where all students belong to both the advanced mathematics and physics populations, all 12 student booklets are distributed among the students according to a predetermined random order, so that approximately equal proportions of students respond to each booklet. In classes containing

only students from the advanced mathematics population, only the six advanced mathematics booklets are distributed. Similarly, in classes with physics students, only the physics booklets are distributed.

Question Types and Scoring Procedures

Students' knowledge and understanding of mathematics and science are assessed through a range of questions in each subject. As described in the *TIMSS 2015 Item Writing Guidelines* (Mullis & Martin, 2013), two question (i.e., item) formats are used in the TIMSS Advanced assessments: multiple-choice and constructed-response. At least half of the total number of points represented by all of the items will come from multiple-choice items. Each multiple-choice item is worth one score point. Constructed-response items generally are worth one or two score points, depending on the nature of the task and the skills required to complete the item. In developing assessment items, the choice of item format depends on the mathematics or physics being assessed as well as the format that best enables students to demonstrate their proficiency.

Multiple-choice Items

In TIMSS, multiple-choice items provide students with four response options, of which only one is correct. These items can be used to assess any of the behaviors in the cognitive domains. Multiple-choice items allow valid, reliable, and economical measurement of a wide range of content in a relatively short testing time. However, because they do not allow for students' explanations or supporting statements, these items may be less suitable for assessing students' ability to make more complex interpretations or evaluations. In developing the multiple-choice items, it is important that the questions and response options are written clearly and concisely in order to minimize the reading load of the question. The options that are incorrect are written to be plausible, but not deceptive. For students who may be unfamiliar with this test question format, the instructions given at the beginning of the test include a sample multiple-choice item that illustrates how to select and mark an answer.

Constructed-response Items

For this type of test item students are required to construct a written response, rather than select a response from a set of options. Because these items allow students to provide explanations, support an answer with reasons or numerical evidence, draw diagrams, or display data, constructed-response items are

particularly well-suited for assessing aspects of knowledge and skills that require students to explain phenomena or interpret data based on their background knowledge and experience.

The scoring guide for each constructed-response item describes the essential features of appropriate and complete responses. The guides focus on evidence of the type of behavior the item assesses. They describe evidence of partially correct and completely correct responses. In addition, sample student responses at each level of understanding provide important guidance to those who will be rating the students' responses. In scoring students' responses to constructed-response items, the focus is solely on students' achievement with respect to the topic being assessed, not on their ability to write well. However, students need to communicate in a manner that will be clear to those scoring their responses.

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

Because constructed-response items constitute an important component of the TIMSS Advanced assessment and are an integral part of the measurement of trends, it is very important for scoring guides to be implemented consistently in all countries and in each data collection year. In order to ensure consistent application of the scoring guides for trend items in the 2015 assessment, IEA has archived samples of student responses to the TIMSS Advanced 2008 assessments from each country; these are used in order to train scorers in 2015 and to monitor consistent application for those items appearing in both assessments.

Score Points

In developing the assessment, the aim is to create item blocks that each provide, on average, about 15 score points. Item blocks contain a variety of item types, including multiple-choice items (1 point each) and constructed-response items (1, 2, or more points) that allow for partial as well as full credit. The exact number of score points and the exact distribution of question types per block varies somewhat.

Releasing Assessment Material to the Public

TIMSS Advanced 2015 is the third in the TIMSS Advanced series of studies, and provides data on trends in mathematics and science achievement over a 20-year period, from 1995 through 2008 to 2015. It is envisaged that, in the future, TIMSS Advanced will be administered on the same four-year schedule as TIMSS (i.e., in 2019, 2023, and so on into the future). With each assessment, as the international reports are published, a selection of items are released in order to provide the public with as much information as possible about the nature and contents of the assessment. At the same time, the measurement of trends is safeguarded by keeping secure a substantial proportion of the items. As items are released, new items will be developed to take their place.

According to the TIMSS Advanced 2015 design, 4 of the 9 assessment blocks in each subject will be released when the assessment results for 2015 are published; the remaining 5 will be kept secure for use in later assessments. The released blocks will include one block containing trend items from 1995, one block of trend items from 2008, and two blocks of items used for the first time in 2015. The released items will be replaced with new items before the next survey cycle, in 2019.

Background Questionnaires

An important purpose of TIMSS Advanced is to identify the procedures and practices that are effective in improving students' learning in advanced mathematics and physics. In order to better understand the contextual factors detailed in Chapter 3 that affect students' learning, TIMSS Advanced administers background questionnaires to students, their teachers, and their school principals. TIMSS Advanced also administers curriculum questionnaires to specialists in order to collect information about educational policies and the national contexts that shape the content and implementation of the advanced mathematics and physics curricula across countries. Finally, the *TIMSS Encyclopedia* provides a more qualitative description of mathematics and science education in the participating countries.

Student Questionnaire

A student questionnaire is completed by each student who takes the TIMSS Advanced assessment. This questionnaire asks about aspects of students' home and school lives, including basic demographic information, their home environment, school climate for learning, and self-perception and attitudes toward advanced mathematics and/or physics. The student questionnaire requires about 30 minutes to complete.

Teacher Questionnaires

A teacher questionnaire is completed by the teachers of the advanced mathematics and/or physics classes sampled to take part in the TIMSS Advanced testing. This questionnaire is designed to gather information on teacher characteristics, the classroom contexts for teaching and learning advanced mathematics and physics, and the topics taught in these subjects.

In particular, the teacher questionnaire asks about teachers' backgrounds, their views on opportunities for collaboration with other teachers, their job satisfaction, and their education and training, as well as professional development. The questionnaire also collects information on characteristics of the classes tested in TIMSS Advanced, instructional time, materials, and activities for teaching mathematics and science and promoting students' interest in the subjects, use of computers, assessment practices, and homework.

Although the general background questions are parallel across advanced mathematics and physics versions of the teacher questionnaire, questions pertaining to instructional and assessment practices, content coverage, and teachers' views about teaching the subject matter are tailored toward advanced mathematics or physics. Many questions, such as those related to classroom activities, are specific to the classes sampled for TIMSS Advanced. This questionnaire requires about 30 minutes of teachers' time to complete.

School Questionnaire

The principal of each school participating in TIMSS Advanced is asked to respond to this questionnaire. It asks about school characteristics, instructional time, resources and technology, parental involvement, school climate for learning, teaching staff, and the role of the principal. This questionnaire is designed to take about 30 minutes to complete.

Curriculum Questionnaires

The National Research Coordinator in each country is responsible for completing the advanced mathematics and physics curriculum questionnaire, drawing on the expertise of curriculum specialists and educators. The questionnaire is designed to collect basic information about the organization of the advanced mathematics and physics curriculum in each country, and about the content of these subjects intended to be covered by the end of secondary schooling.

TIMSS 2015 Encyclopedia

The *TIMSS 2015 Encyclopedia* provides context for mathematics and science instruction in the participating countries. Countries participating in TIMSS 2015 at the fourth and/or eighth grades each contribute a chapter to the encyclopedia. Those countries participating in TIMSS Advanced 2015 also will include information about advanced mathematics and physics curricula and instruction.