TIMSS 2015

ITEM WRITING GUIDELINES

Ina V.S. Mullis and Michael O. Martin
Executive Directors

TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

Table of Contents

Summary of TIMSS 2015 Item Writing Process and Guidelines 1

Introduction 2

General Issues in Writing Items for the TIMSS 2015 Assessments 2

Writing Multiple-choice Items 9

Writing Constructed-response Items and Scoring Guides 14

Documenting the TIMSS 2015 Items..... 23

Reviewing Items and Scoring Guides 23

Appendix A: Multiple-choice Item Review Checklist 26

Appendix B: Constructed-response Item and Scoring Guide Review Checklist 27

References..... 28

Summary of TIMSS 2015 Item Writing Process and Guidelines

Typically, participants will work in groups of two or three. Each group will be assigned specific content areas. Participants will be writing items in English and saving them as Microsoft Word files that will be collected at the end of each day.

When writing items, PLEASE:

1. Address the TIMSS 2015 or TIMSS Advanced 2015 Assessment Frameworks. Write questions that match the topics in each content domain, and pay particular attention to writing questions that cover the range of the three cognitive domains. In accordance with the frameworks, write questions that address the applying and reasoning domains, as well as the knowing domain.
2. Consider the best item format for the question. About half of the items you develop should be multiple-choice and the other half should be constructed-response items worth 1 or 2 score points.
3. For each item, consider the timing, grade appropriateness, difficulty level, potential sources of bias (cultural, gender, or geographical), and ease of translation. Make sure that item validity is not affected by factors that unnecessarily increase the difficulty of the item, such as unfamiliar or overly difficult vocabulary, grammar, directions, contexts, or stimulus materials.
4. For multiple-choice items, keep the guidelines for writing multiple-choice questions in mind. In particular—ask a direct question, make sure there is one and only one correct answer, and provide plausible distracters.
5. For constructed-response questions, write a full-credit answer to the question in terms of the language, knowledge, and skills that a student in the target grade could be expected to possess. This tests the clarity of the question and also provides guidance about whether to allocate 1 or 2 score points to the item.
6. Develop a specific scoring guide for each constructed-response item.

Introduction

These guidelines are to help ensure that the best possible items are developed for TIMSS 2015, TIMSS Numeracy 2015, and TIMSS Advanced 2015. The TIMSS & PIRLS International Study Center has developed these guidelines for writing and reviewing items and scoring guides to facilitate successful item development. It is important to follow some basic procedures so that the TIMSS assessments are uniform in approach and format. During the item-writing sessions, please ask staff or consult these guidelines if you have any questions.

General Issues in Writing Items for the TIMSS 2015 Assessments

Item writing is a task that requires imagination and creativity, but at the same time demands considerable discipline in working within the assessment frameworks and following the guidelines for item construction provided in this manual. These guidelines pertain to good item and test development practices in general, and have been collected from a number of sources. They are designed to help produce items that measure achievement in mathematics and science fairly and reliably, and that enhance the validity of the TIMSS assessments. All of the following issues must be considered in judging the quality and suitability of an item for inclusion in the TIMSS field tests.

Alignment with the Frameworks

Consistent with the principles of evidence-centered design (e.g., Mislevy, Almond, & Lukas, 2003), the TIMSS 2015 assessment is based on:

- Detailed content and cognitive domain descriptions organized into frameworks for each assessment;
- Items aligned with the content topics and cognitive domains and designed to collect evidence about what students know and are able to do; and

- Scoring guides with well-defined categories and a detailed description of the kind of responses that belong in each category.

The TIMSS assessment frameworks in mathematics and science describe those outcomes generally regarded as important at the fourth and eighth grades; the TIMSS Advanced assessment frameworks in advanced mathematics and physics describe those outcomes generally regarded as important for students preparing to enter STEM careers. It is fundamental that every item written for mathematics or science measures the two things described in the TIMSS or TIMSS Advanced 2015 frameworks:

- One of the **content topics** (mathematics) or **performance objectives** (science), and
- One of the **cognitive domains**.

When preparing to produce an item for fourth grade, eighth grade, or Advanced, the first step is to focus on the content topic or performance objective to be assessed. When writing each item, remember that it also contributes to a measure of proficiency in a cognitive domain. These two elements together provide evidence about what students know and are able to do. Keep in mind that TIMSS 2015 assesses student learning in particular topics. Think:

- What should the student know?
- What should the student be able to do?
- What kind of evidence best demonstrates this knowledge or ability?

That is, what knowledge does this item allow a student to show? What cognitive processes does this item require a student to demonstrate? What task best allows the student to demonstrate this knowledge or ability?

The TIMSS 2015 science framework and the TIMSS Advanced 2015 physics framework include sections on science practices fundamental to all science disciplines. Some of the items developed for TIMSS 2015 fourth- and eighth-grade science and TIMSS Advanced 2015 physics should not only focus on a specific performance objective and cognitive

domain, but should also produce evidence that a student can employ skills associated with practicing science.

Types of Items

TIMSS includes two types of items—multiple-choice items where the student chooses the correct answer from four response options, and constructed-response items where the student is required to provide a written response.

PLEASE keep item format in mind. About half of the items you develop should be multiple-choice and half should be constructed-response.

- **Multiple-choice items** allow valid, reliable, and economical measurement of a wide range of content in a relatively short testing time.
- **Constructed-response items** allow students to provide explanations, support an answer with reasons or numerical evidence, draw diagrams, or display data.

If you think of another item type, it may be used as long as it provides valid measurement and is feasible to administer and to score reliably.

Testing Time

When developing items, it is important to consider the time required for students to complete the required task. The amount of time required to complete an item should be consistent with the time allotment for items in the overall test design. As a general rule, a multiple-choice item on TIMSS 2015 is expected to require about 1 minute or less to complete, and constructed-response items are allocated 1-3 minutes. On average, a TIMSS Advanced 2015 item should require 3 minutes. Items should be designed to require the appropriate amount of time.

Grade-Appropriate Language and Context

In keeping with the principles of universal design (e.g., Dolan & Hall, 2007) for assessment items and tasks, the language, style, and reading level used in items must be accessible to a range of students in the target grades. Keep the language as simple as possible, and take care to

use grade-appropriate vocabulary and terms. The reading level of items should be at an elementary level for the target grade. In general, the amount of reading should be kept to a minimum, given the context of the problem. Write questions in the active voice (i.e., doer of action (subject) before action (verb)) and avoid conditional words, clauses, and tenses (e.g., if, suppose, when).

The context for the item may relate only to the discipline of mathematics or science, or to aspects of those subjects encountered in everyday life. However, if the item involves a "real-world" setting make sure the setting is familiar to students. Avoid using context-specific vocabulary that may not be familiar to all students. An unnecessarily complicated item context or unfamiliar context-specific vocabulary may artificially increase the difficulty of the item and pose a threat to item validity.

Item Difficulty

Information from individual TIMSS 2015 items should provide valuable insight into student learning by providing evidence about what the student knows or is able to do. Additionally, each of the items needs to contribute to the overall mathematics or science assessment. It is desirable that there be some relatively easy items and some challenging items. However, items that almost all students or almost no students are able to answer correctly reduce the effectiveness of the test to discriminate between groups with high achievement and groups with low achievement. Typically, the majority of items used in the final test will be ones that are answered correctly by 30 to 70 percent of the students on average internationally.

Avoiding Bias

When preparing assessment items, be sensitive to the possibility of unintentionally placing particular groups of students at an unfair disadvantage. An international study requires special attention to the diversity of environments, backgrounds, beliefs, and cultures among students in the participating countries.

Considering National Contexts

Be particularly aware of issues related to nationality, culture, ethnicity, and geographic location. Items requiring background knowledge confined to a subset of participating countries are unlikely to be suitable.

Geographic location has an effect on students' learning experiences, as aspects of the local environment have an impact on schooling. Even though television and the Internet can provide students with some knowledge of remote places, firsthand experience of some phenomena enhances understanding and can give some students an advantage over others.

Gender

A gender-related context included in an item may distract some students from the purpose of the item. Situations in which stereotypical roles or attitudes are unnecessarily attributed to males or females, or in which there is implicit disparagement of either gender, are not acceptable.

Facilitating Comparable Translation

The international version of items will be in American English. Therefore, items developed at this meeting must be submitted in English. Keep in mind, however, that after review and revision, the items selected for the field test and main data collection will be translated from English into the languages of instruction of the countries in the study.

Please be sensitive to issues that might affect how well items can be translated to produce internationally comparable items. The TIMSS 2015 translation procedures do allow names and places to be changed to what is appropriate for a country, provided the essential nature and difficulty of the item are not altered. Idioms and expressions that defy translation must be avoided.

Problems Involving Money

Problems involving computations with money, especially those set in "real life" contexts, are problematic for international studies. The cost of

a common article in one country may be a fraction of the base unit of currency, while the same article in another country may cost thousands of the base unit. In some countries, the cost of an article may never include a decimal point. If the inclusion of costs is an essential part of a problem, use "Zeds". This is the TIMSS fictitious unit of currency, which enables each country to work with the same numbers.

Graphics

Take special care to ensure that diagrams and graphs are drawn accurately (to scale unless otherwise noted), and are correctly and fully labeled. Any graphics included in an item should be necessary in order to solve the problem or to answer the question and should be adequately explained and referred to directly within the item, as indicated by the principles of universal design for assessment items. In particular, visual elements should:

- Align with the wording and task presented in the item text;
- Depict only information necessary to solve the problem or answer the question so as not to distract or confuse students;
- Be included to emphasize an important part of an item if their inclusion makes the item accessible for more students; and
- Be labeled clearly.

Graphics for items may be submitted as a hand-drawn paper document.

All graphics images must be able to be viewed equally well on a computer screen and on a printed page. In particular,

- When using color or greyscale, choose images with few colors and a limited amount of shading.
- Do not reference specific colors in item prompts (e.g., the blue line on the graph represents…).

Copyright

All of the items developed for the TIMSS 2015 assessments, will be copyrighted by the IEA. For this copyright to be valid, it is important that TIMSS items do not infringe on other copyrights. All of the items used in TIMSS 2015 must be specifically developed for TIMSS 2015 and never used in other assessments. Also, in developing items for

TIMSS 2015 any copyrighted stimulus material must be acknowledged appropriately. For example, statistical graphs from publications or extracts from articles in publications that are used in an item must be identified appropriately, and full details about the sources must be submitted with the item.

Pattern Items—TIMSS 2015 Mathematics ONLY

In developing items assessing the topics: Number–Fourth Grade, "Identify and use relationships in a well-defined pattern (e.g., describe the relationship between adjacent terms and generate pairs of whole numbers given a rule)," and Algebra–Eighth Grade, "Generalize pattern relationships in a sequence, or between adjacent terms, or between the sequence number of the term and the term, using numbers, words, or algebraic expressions," the patterns must be well defined in the question. The question needs to describe that the pattern "repeats every four shapes" or "increases by the same amount from one number to the next." For example, the question might state: "Ellen made a number pattern using the rule add 4." For geometric pattern items, it is possible to say "the same rule is used to get from one figure to the next" (if it is the same, e.g., adding 2 circles and a square to each figure). Often the algebra pattern items should be in the constructed-response format asking students to justify or explain the rule for the pattern.

Use of Calculators—TIMSS 2015 Eighth Grade and TIMSS Advanced 2015 ONLY

Students participating in the TIMSS 2015 eighth grade and TIMSS Advanced 2015 assessments will be permitted to use calculators for the entire assessment. Keep in mind that today's calculator technology is quite advanced and some, but not all, calculators are capable of not only graphing tasks, but also symbolic as well as numerical algebra and calculus tasks. Every effort should be made to ensure that the items do not advantage or disadvantage students either way—with or without calculators. Calculators are not permitted at the fourth grade.

Formula Sheets—TIMSS Advanced 2015 ONLY

The booklets containing TIMSS Advanced 2015 mathematics and physics items will also contain pages with information about advanced

mathematics and physics notation, selected formulas from advanced mathematics and physics, and selected physics constants. When writing items for which a formula is necessary to solve the problem or answer the question, please also include the formula(s) with the distracter analysis (multiple-choice items) or the scoring guide (constructed-response items).

Writing Multiple-choice Items

A multiple-choice item asks a question or establishes the situation for a response. For the TIMSS 2015 assessments, this type of item includes four response choices, or options, from which the correct answer is selected. A multiple-choice item is characterized by the following components:

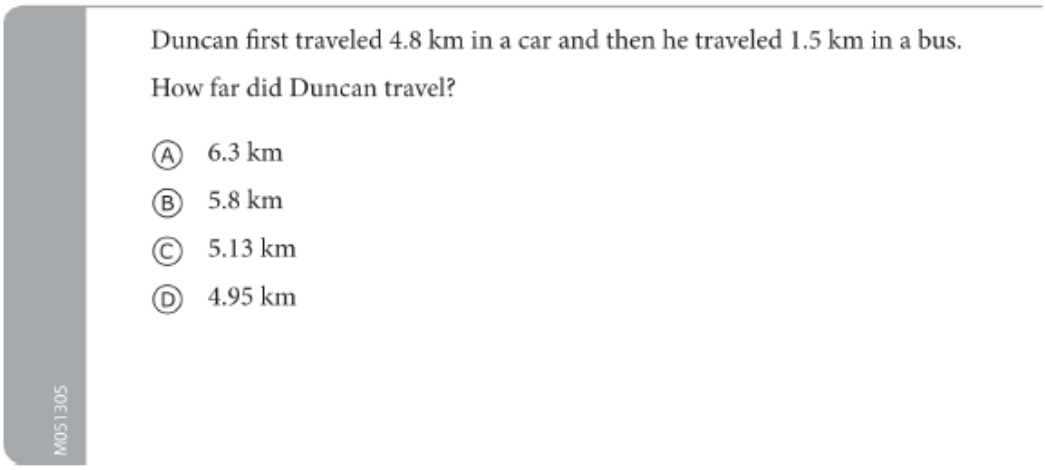
- The stimulus presents the contextual information relevant to the item.
- The stem presents the question or prompt the student must answer.
- The options refer to the entire set of labeled response choices presented under the stem.
- The key is the correct response option.
- The distracters are the incorrect response options.

At least half of the items developed for TIMSS 2015 will be multiple-choice items. The next sections present guidelines specific to multiple-choice items, and include information about writing the stem, structuring the response options, developing plausible distracters, and providing a distracter analysis.

PLEASE keep the guidelines for writing multiple-choice questions in mind. In particular, ask a direct question, make sure there is one and only one correct answer, and provide plausible distracters.

The Stem

For the TIMSS 2015 assessments, since clarity is of vital importance, please phrase all stems as a **direct question**. The following is an example of a stem formulated as a question:



1. Provide sufficient information in the stem to make the question clear and unambiguous to students. In nearly all cases, the question must be able to stand alone, and be answerable without the response options. An exception would be items asking students to choose the best estimate of a quantity.
2. The stem should not include extraneous information. Extraneous information is liable to confuse students who might otherwise have determined the correct answer.
3. Avoid using negative stems—those containing words such as NOT, LEAST, WORST, EXCEPT, etc. If it is absolutely necessary to use a negative stem, highlight the negative word, (e.g., capitalize, underline, or put in bold type so that it stands out for the student). If the stem is negative, use only positive response options—do not use double negatives.
4. If there is not one universally agreed upon answer to the question, it is best to include "of the following" or some similar qualifying phrase in the stem.

5. Avoid questions for which a wrong method yields the correct answer (e.g., a question about a circle with a radius of 2; because $2r = r^2$, students computing either the area or the circumference get 4π).

Structure of the Response Options (or Alternatives)

1. Write multiple-choice items with four response options, labeled A–D (as shown in the example item about distance traveled, above).
2. Make sure that one of the four response options or alternatives is the key or correct answer. Make sure there is only one correct or best answer.
3. Make sure that the four response options are independent. For example, response options should not represent subsets of other options. Also, do not include pairs of response options that constitute an inclusive set of circumstances (e.g., day or night, does or does not).
4. Make sure that the grammatical structure of all response options "fit" the stem. Inconsistent grammar can provide clues to the key or eliminate incorrect response options. Avoid writing items where the options complete a sentence begun in the stem, because these can cause problems with translation.
5. Make sure all (or sets) of the response options are parallel in length, level of complexity, and grammatical structure. Avoid the tendency to include more details or qualifications in the correct response, thus making it stand out. If the options are not parallel in length, please order the options short to long if at all possible.
6. Do not use words or phrases in the stem that are repeated in one of the response options and, therefore, act as a clue to the correct response.
7. Do NOT use "none of these" and "all of these" as response options.
8. Arrange the response options in a logical order if this makes sense and saves the student time in reading the options (e.g., years in chronological order, numbers from least to greatest).

9. Avoid writing items where students can work backwards from the response options to find the correct answer (e.g., solving for *x* in an equation). Sometimes described as "plug and chug" items, such questions or problems will not be included in any of the TIMSS 2015 assessments. In such cases, a constructed-response item may be more appropriate than a multiple-choice item.

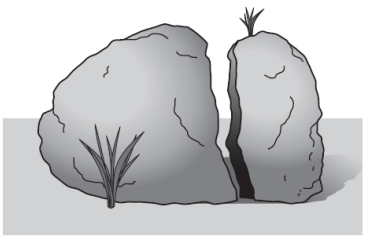
Plausibility of Distracters

Use plausible distracters (incorrect response options) that are based on likely student errors or misconceptions. This reduces the likelihood of students arriving at the correct response by eliminating other choices and, equally important, may allow identification of widespread student misunderstandings or tendencies that could lead to curricular or instructional improvements. If there are no plausible errors or misconceptions, still make the options "reasonable." For example, they should be from the same area of content. However, avoid the use of "trick" distracters.

Distracter Analysis

Please include a brief analysis of each response option or rationale for inclusion of specific response options with your item (one sentence at the most for each response option). For example:

Scientists think that the rocks in the picture were once a single rock.



Which property of water had the **most** effect on splitting the rock into two pieces?

Ⓐ Water expanding when it freezes.

Ⓑ Water boiling at 100°C.

Ⓒ Water having a density less than rock.

Ⓓ Water dissolving many substances.

S052254

Distracter rationale:

A. [Key]

B. Assumes that boiling water heats up the rock and results in its separation into two pieces

C. Associates the density difference between water and rock with water acting to split the rock.

D. Assumes that water dissolves the rock in such a way that the two pieces result.

Writing Constructed-response Items and Scoring Guides

For some desired outcomes of mathematics and science education, constructed-response items provide more valid measures of achievement than do multiple-choice items. The quality of constructed-response items depends largely on the ability of scorers to assign scores consistently and reliably within and across countries. Thus, it is essential that each constructed-response item and its scoring guide be developed together.

PLEASE keep the guidelines for writing constructed-response questions in mind. In particular, ask a clear question, and develop a scoring guide for the question at the same time as the question is developed.

Constructed-response items usually require students to give a numerical result, provide a short explanation or description given in one or two phrases or sentences, complete a table, or provide a sketch. They are scored as either 1 or 2 points for fully-correct answers.

- 1-point constructed-response items are scored as correct (1 score point) or incorrect (0 score points).
- 2-point constructed-response items are scored as fully correct (2 score points), partially correct (1 score point), or incorrect (0 score points). For example, a response demonstrating thorough understanding of concepts and processes will receive full credit (2-points). These responses show a complete or deeper understanding than a response that will receive partial credit (1-point). (Developing scoring guides is explained in the next section.)

Constructed-response items should be used when it is desirable that the student be required to think of an answer without the possible cues provided by an option in a multiple-choice item. If too few plausible distracters are available for a multiple-choice item, it may be better framed as a constructed-response item.

Developing a constructed-response item accurately targeted on the ability to be assessed, along with the accompanying scoring guide, is not a straightforward task. Care in writing constructed-response items is especially important for two reasons. First, if the task is not well specified students may interpret the task in different ways and respond to different questions. Second, a constructed-response item may carry more score points than a multiple-choice item.

Guidelines for Writing Constructed-response Items

1. Students will not be allowed to ask the test administrator for clarification. Write questions in easily accessible language appropriate to the age and experience of the target population. Use simple vocabulary and sentence structure, and avoid using complicated names for the subjects in the item.
2. Make what is expected of students as clear as possible without compromising the intent of the item. Give an indication, where appropriate, of the extent, or level of detail, of the expected answer (e.g., "Give three reasons …" rather than "Give some reasons …" and "Draw a labeled diagram illustrating the water cycle" rather than "What is meant by the term 'water cycle'?"). Select real life problem settings that are likely to be "real" to students at the target grade levels, and that involve quantities that are realistic for the situations.
3. Avoid asking questions that could give rise to answers that cannot be scored strictly in terms of accuracy of mathematical or scientific understanding (e.g., "What are satellites used for?").
4. Students should be able to complete the task in the time allocated for each constructed-response item, that is, a maximum of 3 minutes.
5. Write an appropriate answer to the question in terms of the language, knowledge, and skills that a good student at the target grade could be expected to possess. This tests the clarity of the question and is also an essential first step in producing a scoring guide for the item. It is also helpful for those who are reviewing the question.

6. Produce a scoring guide (see below). This action usually results in amendments to the item to clarify its purpose and improve the quality of information that can be obtained from student responses.

Writing Scoring Guides

To ensure reliability, constructed-response items need scoring guides with well-defined categories for allocating score points. It also is important to collect information of value for educational improvement. Students' answers can provide insights into what they know and are able to do, including common misconceptions.

The TIMSS Generalized Scoring Guidelines

The generalized scoring guidelines used for 1- and 2-point constructed-response items are described in Table 1.

Table 1: TIMSS Generalized Scoring Guidelines for Constructed-response Items

Score Points for 1-point system
1 Point (Full credit) A one-point response is correct. The response indicates that the student has completed the task correctly.
0 Points (No credit) A zero-point response is incorrect, irrelevant, or incoherent.
Score Points for 2-point system
2 Points (Full credit) A two-point response is complete and correct. The response demonstrates a thorough understanding of the concepts and/or procedures embodied in the task. • Indicates that the student has completed all aspects of the task, showing correct application of concepts and/or procedures • Contains clear, complete explanations, supporting work, or evidence when required
1 Point (Partial credit) A one-point response is only partially correct. The response demonstrates only a partial understanding of the concepts and/or procedures embodied in the task. • Addresses some elements of the task correctly but may be incomplete • May contain a correct answer but an incomplete explanation when required • May contain an incorrect answer with an explanation or supporting work indicating a correct understanding of the concepts
0 Points (No credit) A zero-point response is inaccurate or inadequate, irrelevant, or incoherent.

The TIMSS Two-digit Diagnostic Scoring System

The TIMSS diagnostic scoring system uses two digits. For example, 10, 11, or 20.

The **first digit** is the score indicating the degree of correctness of the response as described in the generalized scoring guidelines.

The **second digit** is used to classify the method used in solving a problem, or perhaps to track common errors or misconceptions. The information from the second digit addresses questions such as: Do approaches that lead to correct responses to the item vary across countries? Is there one approach that students have more success with than others? What are the common misconceptions that students have about the matter being tested? What common errors are made?

The First Digit

The **first digit for correct or partially correct responses** signifies the number of score points given to the response. Thus:

The first digit for correct responses is 1 for one point or 2 for two points. When TIMSS started in the early 1990s, it was decided not to use 0 for the first digit. Thus:

- The first digit for incorrect responses is 7.
- The first digit for a blank response is 9.

The Second Digit

The **second digit for correct or incorrect responses** provides diagnostic information. Thus:

- The second digits used for diagnostic purposes with either correct or incorrect responses can be 0 through 2 (codes 20–22, 10–12, and 70–72).

- However, it is unusual for an item to give rise to more than two commonly used correct methods, or more than one common error or misconception. Frequently no more than one or two categories are required. In other words, the specific diagnostic codes should capture only the predominant correct and incorrect approaches/strategies used by students. Scoring of constructed-response items is a significant cost factor for national centers, so care should be taken not to provide codes for response types that do not have apparent value for educational improvement.
- Since not all incorrect student responses should be categorized into pre-defined categories, for codes with a first digit of 7, the second digit of 9 is used to designate a response that is "other" than any specific diagnostic codes included in the guide. Thus, an incorrect response not fitting a pre-defined incorrect code is given a 79 for "other incorrect." If no diagnostic categories are defined, all incorrect responses are coded 79.
- **Code 99** means a completely **BLANK** response.

Examples of Scoring Guides

The following examples are given to illustrate the diagnostic scoring guides used in TIMSS (grade 4 and grade 8) and TIMSS Advanced (advanced mathematics and physics).

Grade 4 Mathematics Item (1 point):

Cooney has to form figures 1 to 4 with matches.

Figures 1, 2, and 3 are shown below.

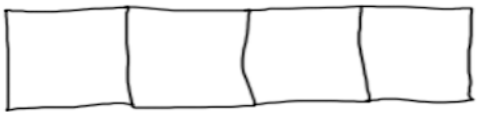
He needs four matches to form figure 1, seven matches to form figure 2, and ten matches to form figure 3.

He uses the same rule each time to make the next figure in the pattern.

1 2 3

How many matches will he need to form figure 4?

Answer: _____

Code	Response	Item: M051601
	Correct Response	
10	13	
	Incorrect Response	
70		
79	Incorrect (including crossed out, erased, stray marks, illegible, or off task)	
	Non response	
99	Blank	

Grade 8 Science Item (2 points):

Diagram 1 Diagram 2

Ice-cold water was placed in a glass pitcher on a hot day (Diagram 1).
Soon afterwards, liquid appeared on the outside of the pitcher (Diagram 2).

Describe the process that caused the liquid to appear on the outside of the pitcher.

Code	Response	Item: S042404
	Correct Response	
20	Describes the process of condensation by referring to water vapor (in the air) condensing on the cool outside surface of the pitcher. Examples: The water droplets came from the water vapor in the air which condenses into liquid water when it touches a cool surface. The surface of the glass pitcher is cool because it loses heat to the ice cold water. It came from the water vapor condensing on the cool surface of a glass pitcher.	
	Partially Correct Response	
10	Describes the process of condensation by referring to water vapor (in the air) condensing without mentioning the coolness of the pitcher. Examples: The liquid came from water vapor condensing.	
11	States condensation without referring to water vapor. Examples: Condensation. It condensed from the air.	
	Incorrect Response	
79	Incorrect (including crossed out, erased, stray marks, illegible, or off task) Examples: Liquid came from the sky. It came from the clouds.	
	Nonresponse	
99	Blank	

Advanced Mathematics Item (1 point):

M3_07

A regular polygon of n sides is inscribed in a circle of radius 1.

What is the value of the limit of the perimeter of the polygon as the number of sides n increases to infinity?

Code	Response	Item: MA13027
	Correct Response	
20	Any of 2 pi, 2π, 6.28, 6.3, or $2\pi = 6.28$	
	Partially Correct Response	
10	$\lim_{n\to\infty} 2n\sin\frac{\pi}{n}$ **Note:** Accept also $\lim_{n\to\infty} 2n\sin\frac{180}{n}$	
11	2 pi r or $2\pi r$ or makes a statement such as "The value of the limit is equal to the circumference of the circle."	
	Incorrect Response	
70	π or pi or 3.14	
71	∞ or "infinity" or "the limit does not exist" or equivalent statement	
79	Other incorrect (including crossed out, erased, stray marks, illegible, or off task) *Examples:* *1.* $\lim_{n\to\infty} 2n\sin\frac{2\pi}{n}$ *or* $\lim_{n\to\infty} 2n\cos\frac{\pi}{n}$ *or similar formula containing error* *2. 1* *3. "Almost a circle"or similar answers in words, not numerical values, stating that the shape of the polygon will become very close to that of a circle.*	
	Nonresponse	
99	Blank	

Physics Item (2 points):

5

The planet Venus, like Earth, revolves around the sun in approximately a circular orbit. Venus is closer to the sun than the Earth is.

Using Newton's Second Law and Law of Gravity, show that Venus moves faster than Earth in its orbit.

Code	Response	Item: PA23022
	Correct Response	
20	A response that includes the following steps 1. States the two laws in mathematical form Newton's Second Law: $F = ma$ and the Law of Gravity: $F = \frac{GMm}{r^2}$ 2. Applies the formula for centripetal acceleration: $a = \frac{v^2}{r}$, 3. Derives the formula for velocity, $v = \sqrt{\frac{GM}{r}}$ (or equivalent) and uses this to show that v (Venus) is greater than v (Earth).	
	Partially Correct Response	
10	Step 1 and 2 complete but not Step 3	
	Incorrect Response	
70	Step 1 only complete.	
79	Other incorrect (including crossed out, erased, stray marks, illegible, or off task)	
	Nonresponse	
99	Blank	

Documenting the TIMSS 2015 Items

During the item-writing sessions, teams will be writing items on computers using Microsoft Word. At the end of each day, the TIMSS & PIRLS International Study Center staff will collect the files from each team.

When writing the TIMSS 2015 items, please use the template that has been provided and complete the necessary documentation as described below.

Filename: Subject (mathematics or science), grade (4, 8, A), and team number (to be assigned)

For each individual item, provide:

1. The TIMSS, TIMSS Numeracy, or TIMSS Advanced 2015 Content Domain, topic area, and topic (or objective for science) the item measures;
2. The TIMSS, TIMSS Numeracy, or TIMSS Advanced 2015 Cognitive Domain and sub-area the item addresses;
3. The item number (1, 2, 3, etc.);
4. The key and distracter analysis (multiple-choice items only); or
5. The scoring guide.

Reviewing Items and Scoring Guides

Items selected for inclusion in the TIMSS 2015 field tests or main data collection will go through a thorough review process involving the TIMSS & PIRLS International Study Center staff, the mathematics and science consultants, the Science and Mathematics Item Review Committee (SMIRC), and the National Research Coordinators. The first step in this item review process begins with you. Item writers are expected to review and revise their own items in accordance with the procedures outlined here and presented in the item-writing sessions. In addition, depending on the time available, the items will be reviewed by other item writing teams.

If it happens that items are written after the NRC meeting, the item writers are expected to arrange to have their items reviewed by at least one independent reviewer in their own country. Any concerns with items and/or scoring guides detected in the course of this review should be corrected prior to submitting items to the TIMSS & PIRLS International Study Center.

Item writers and item reviewers must be very critical when reviewing items and the item writers should expect to have to explain their items. The earlier necessary changes are made to items, the better. Last minute changes to items to remove errors often result in other flaws being introduced.

The following sections provide guidelines for the review of multiple-choice items and constructed-response items together with their scoring guides and are to be used by item writers and reviewers. To facilitate item review, item review checklists for multiple-choice and constructed-response items are provided in Appendix A and Appendix B, respectively, of this manual.

Reviewing Multiple-choice Items

In reviewing each multiple-choice item, item reviewers should:

1. Identify what they consider to be the (only) correct response and compare this with that originally identified by the item writer.
2. Check that their judgments of the TIMSS 2015 content and cognitive classifications correspond with those indicated by the item-writing team.
3. Check the item against each of the entries in the Multiple-choice Item Review Checklist (see Appendix A).
4. Identify and note any concerns with the item.

Reviewing Constructed-response Items

In reviewing each constructed-response item, item reviewers should:

1. Check that their judgments of the TIMSS 2015 content and cognitive classification correspond with those indicated by the item-writing team.
2. Check the item against each of the entries in the Constructed-response Item and Scoring Guide Review Checklist (See Appendix B).
3. Write an outline of what s/he believes would be a good response to the item for a student at the target grade. Review the scoring guide for the item, comparing it with your response, to make sure that you agree with the number of score points allocated and the clarity of the distinction made between the levels. Also, see if the most likely types of student responses have been categorized.
4. Identify and note any concerns with the item and/or scoring guide.

Appendix A: Multiple-choice Item Review Checklist

Item Characteristic	Yes	No
Is the mathematics/science correct?		
Task clear to students?		
Free of cultural, gender, or geographical bias?		
Seems to be OK for translation?		
Negative stem avoided (or negative word highlighted if used)?		
One (only) correct response?		
Distracters plausible but demonstrably incorrect?		
Options parallel in structure?		
Words in stem NOT repeated in options?		
Content classification correct?		
Cognitive classification correct?		

Appendix B: Constructed-response Item and Scoring Guide Review Checklist

Item Characteristic	Yes	No
Is the mathematics/science correct?		
Task clear to students?		
Free of cultural, gender, or geographical bias?		
Seems to be OK for translation?		
No unfamiliar factors contributing to difficulty?		
Clear expectations for full-credit response?		
Task can be completed in a reasonable time?		
Scoring guide has appropriate correct and incorrect categories?		
Scoring guide has appropriate number of score points?		
Scoring guide descriptors clear?		
Content classification correct?		
Cognitive classification correct?		

References

Dolan, R.P., & Hall, T.E. (2007). Developing accessible tests with universal design and digital technologies: Ensuring we standardize the right things. In L. L. Cook & C. C. Cahalan (Eds.), *Large-scale assessment and accommodations: What works* (pp. 95-111). Arlington, VA: Council for Exception Children.

Mislevy, R.J., Almond, R.G., & Lukas, J.F. (2003). *A brief introduction to evidence-centered design*. ETS (Research Report RR-03-16). Princeton, NJ: Educational Testing Service.

BOSTON
COLLEGE

timss.bc.edu

International Association
for the Evaluation of
Educational Achievement