AMERICAN INSTITUTES FOR RESEARCH®

MAY 2016

# A Comparison Study of the Program for International Student Assessment (PISA) 2012 and the National Assessment of Educational Progress (NAEP) 2013 Mathematics Assessments

## AIR-NAEP Working paper #02-2016


Kim Gattis
Young Yee Kim
Maria Stephens
Linda Dager Hall
Fei Liu
Juliet Holmes



The research contained in this working paper was commissioned by the National Center for Education Statistics (NCES). It was conducted by the American Institutes for Research (AIR) in the framework of the Education Statistics Services Institute Network (ESSIN) Task Order 14: Assessment Division Support (Contract No. ED-IES-12-D-0002/0004) which supports NCES with expert advice and technical assistance on issues related to the National Assessment of Educational Progress (NAEP). AIR is responsible for any error that this report may contain. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

# A Comparison Study of the Program for International Student Assessment (PISA) 2012 and the National Assessment of Educational Progress (NAEP) 2013 Mathematics Assessments

## AIR-NAEP Working paper #02-2016

**May 2016**

**Kim Gattis**
**Young Yee Kim**
**Maria Stephens**
**Linda Dager Hall**
**Fei Liu**
**Juliet Holmes**

AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

**www.air.org**

# AIR

Established in 1946, with headquarters in Washington, D.C., the American Institutes for Research (AIR) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally in the areas of health, education, and workforce productivity. For more information, visit www.air.org.

# NCES

The National Center for Education Statistics (NCES) is the primary federal entity for collecting and analyzing data related to education in the U.S. and other nations. NCES is located within the U.S. Department of Education and the Institute of Education Sciences. NCES fulfills a Congressional mandate to collect, collate, analyze, and report complete statistics on the condition of American education; conduct and publish reports; and review and report on education activities internationally.

# ESSIN

The Education Statistics Services Institute Network (ESSIN) is a network of companies that provide the National Center for Education Statistics (NCES) with expert advice and technical assistance, for example in areas such as statistical methodology; research, analysis and reporting; and Survey development. This AIR-NAEP working paper is based on research conducted under the Research, Analysis and Psychometric Support sub-component of ESSIN Task Order 14 for which AIR is the prime contractor. The two other sub-components under Task 14 are Assessment Operations Support and Reporting and Dissemination.

The NCES Project officer for the Research, Analysis and Psychometric Support sub-component of ESSIN Task Order 14 is William Tirre (William.Tirre@ed.gov).

The NCES project officers for this research study are Samantha Burg (Samantha.Burg@ed.gov) and Dana Kelly (Dana.Kelly@ed.gov).

## Suggested citation:

Gattis, K., Kim, Y.Y., Stephens, M., Hall, D.H., Liu, F., & Holmes, J. (2016). A Comparison Study of the Program for International Student Assessment (PISA) 2012 and the National Assessment of Educational Progress (NAEP) 2013 Mathematics Assessments. AIR-NAEP working-paper #02-2016. Washington, DC: American Institutes for Research.

## For inquiries contact details:

***Kim Gattis***, Principal Researcher
Email: kgattis@air.org

***Markus Broer***, Project Director for Research under ESSIN Task 14
Email: mbroer@air.org

***Mary Ann Fox***, Project Director of ESSIN Task 14
Email: mafox@air.org

# Executive Summary

In the United States, nationally representative data on student achievement come primarily from two sources: the National Assessment of Educational Progress (NAEP)—also known as "The Nation's Report Card"—and U.S. participation in international assessments, including the Program for International Student Assessment (PISA). Together, these national and international sources provide important information on the performance of U.S. students in key subjects, such as mathematics, science, and reading. While the national assessment provides data on achievement that is tailored to students' school experiences in the United States, the international assessments allow U.S. student performance to be benchmarked to that of students in other countries.

In the winter of 2013, assessment results in the area of mathematics were released for eighth-graders in NAEP 2013 and the 15-year-olds assessed in PISA 2012. NCES thus commissioned a study to compare the two mathematics assessments so that researchers, educators, the mathematics community, and the public could gain a deeper understanding of the similarities and differences in the mathematics assessed in each program and what each assessment contributes to the knowledge base about U.S. students' mathematics performance.

The major findings from this comparison study are as follows:

## Related to the content of the item pools

- *The PISA item pool is more like the NAEP grade 8 item pool than either the NAEP grade 12 or grade 4 item pool in terms of the mathematical content being assessed*. When PISA items were examined for alignment to the NAEP framework,[1] the majority of items (75 percent) were aligned to objectives in the framework at grade 8, with smaller percentages matching objectives in the framework at grade 12 (14 percent) or grade 4 (7 percent).
- *The percentage distribution of items across content areas and categories differs between PISA and NAEP.* When PISA items were examined in the context of the NAEP framework (and then compared to how the NAEP items are classified in the NAEP framework), some differences in relative emphasis became apparent. At grade 8, whereas the NAEP item pool had a stronger emphasis on algebra, the PISA item pool had a stronger emphasis on number properties and operations, and data analysis, statistics, and probability. At grade 12, NAEP and PISA placed a similar emphasis on algebra and on data analysis, statistics, and probability, but the PISA item pool more strongly emphasized geometry and, in contrast to NAEP, there were no items in number properties and operations.

[1] In this study, all PISA items were compared to the NAEP framework and a subset of NAEP items were compared to the PISA framework, for reasons elaborated in the report. The Executive Summary draws primarily on the comparisons to the NAEP framework, which provided data on all items for both assessments.

- *PISA does not assess some mathematics subtopics that NAEP assesses at grade 8 and grade 12*. The subtopics that are included in the NAEP framework but are not explicitly assessed in PISA are estimation; mathematical reasoning using numbers; position, direction, and coordinate geometry; mathematical reasoning in geometry; measurement in triangles; experiments and samples; mathematical reasoning with data; and mathematical reasoning in algebra.
- *PISA and NAEP differ in the distribution of items across levels of cognitive complexity*. Comparing items across NAEP's levels of complexity (low, moderate, and high) reveals a lower percentage of PISA than NAEP items in the low category (44 percent in PISA compared to 56 percent in NAEP at both grades 8 and 12) and a higher percentage of PISA than NAEP items in the moderate category (53 percent in PISA compared to 39 percent in NAEP at both grades 8 and 12). Neither assessment had very large percentages of items in the high category: 2 percent of PISA items and 5 percent of NAEP grade 8 and grade 12 items.

## Related to the item features

- *PISA and NAEP items differ on the extent to which students are required to mathematize problem situations*—that is, to make mathematical sense of real-world situations. PISA explicitly aims to measure students' ability to mathematize; thus, some of the items on the PISA assessment require students to make assumptions or ignore information that is not relevant to the problem. NAEP items typically do not require such aspects.
- *In PISA, the role of context is to set up the problem situations that require students to mathematize (and thus it is integral to the functioning of the items)*, whereas in NAEP, the context does not typically play this role and thus is not integral to the functioning of items.
- *PISA items often have more text than NAEP items, which indicates a heavier reading load*. Some of the added text in a PISA item (or set of items) is due to the intention to make the context more realistic; thus, information may be included that is irrelevant to answering the question. In some instances, additional text is needed to accommodate a translation issue across languages. In contrast, in NAEP, the reading load is kept to a minimum, and generally all of the information provided in an item is relevant to solving the problem. Such differences in the amount of text do not necessarily imply differences in text complexity or overall reading difficulty, which were not examined in this study.
- *Both PISA and NAEP use a variety of representations in their mathematics items, but there are differences in both the quantity and characteristics of these representations*. In NAEP, representations are more likely to be simplified (e.g., a street map converted into simple lines) or to be confined to a small set of easily recognizable types of graphs (e.g., bar graphs or line graphs). PISA is more likely to include photos, realistic diagrams, and unconventional graphs.

In summary, this study found many similarities between the two assessments. However, it also found important differences in the relative emphasis across content areas or categories, in the role of context, in the level of complexity, in the degree of mathematizing, in the overall amount of text, and in the use of representations in assessments

# Contents

**Page**

# Tables

# Exhibits

**Page**



# Figures

**Page**

# 1. Introduction

In the United States, nationally representative data on student achievement come primarily from two sources: the National Assessment of Educational Progress (NAEP)—also known as "The Nation's Report Card"—and U.S. participation in international assessments, including the Program for International Student Assessment (PISA), the Trends in International Mathematics and Science Study (TIMSS), and the Progress in International Reading Literacy Study (PIRLS). While the international assessments may appear to have significant similarities with NAEP, each was designed to serve a specific purpose and each is based on a separate and unique framework and set of assessment items. Thus, each gives a somewhat different view of U.S. student performance.

In December 2013, the U.S. Department of Education's National Center for Education Statistics (NCES) released U.S. results from PISA 2012. PISA assesses the reading, mathematics, and scientific literacy of 15-year-old students, comparing the performance of U.S. students with that of students in other countries. In 2012, PISA had a special focus on mathematics. In November 2013, NCES also released NAEP results in mathematics for eighth-graders—who are close in age to PISA's students—for 2013.

NCES thus recently commissioned a study to compare the PISA 2012 and NAEP 2013 mathematics assessments. While it builds on the foundation of numerous prior NCES comparison studies (see appendix A), this study also expands the scope and introduces some new methodologies that can provide more in-depth data and allow researchers, educators, the mathematics community, and the public to build a deeper understanding of the mathematics being assessed in each program.

The results of this PISA-NAEP comparison study are described in this report. It examines the similarities and differences between the frameworks that guided the development of the PISA 2012 and NAEP 2013 assessments and the items that were used in the assessments.[2]

The purpose of the study is to provide information that supplements the results and describes what each assessment can uniquely contribute to the knowledge base on student performance in mathematics. Educating stakeholders about the similarities and differences between the programs has been—and continues to be—important in helping avoid misunderstanding or misinterpretation of results.

Three main research questions guided the study:

- *Research question 1 (related to the assessment frameworks)*. How similar or different are PISA 2012 and NAEP 2013 (grade 8 and grade 12) in terms of their approaches to mathematics and what they intend to measure?

[2] In this report, "NAEP framework" refers to both the assessment framework and the assessment specifications document, while "PISA framework" refers only to the assessment framework.

- *Research question 2 (related to the item pools)*. How similar or different are PISA 2012 and NAEP 2013 (grade 8 and grade 12) in terms of (a) the mathematics knowledge and skills assessed and (b) the role that context plays in the functioning of the mathematics items?
- *Research question 3 (related to the item features)*. How similar or different are the features (e.g., reading load, use of representations) of the PISA 2012 and NAEP 2013 (grade 8 and grade 12) mathematics items?

The study was conducted by research staff at the American Institutes for Research (AIR) in conjunction with two leading mathematics experts. Two key AIR research staff and the two external experts convened in a series of multi-day meetings during June 2013 to investigate the research questions.[3]

Following this introductory section, the report presents an overview that describes the basic features and design of the PISA 2012 and NAEP 2013 mathematics assessments. The overview is followed by three sections describing the specific methodology and results for the framework comparisons (research question 1), the comparison of the item pools to the framework-related dimensions (research question 2), and the in-depth review of the item features (research question 3), respectively. The report ends with a conclusion that summarizes the key findings from the study.

# 2. Overview of Key Features and Design of PISA and NAEP Mathematics Assessments

This section provides an overview of the PISA and NAEP mathematics assessments, focusing on the unique development of each assessment program, the differences and similarities in their underlying approaches to assessing mathematics, the target populations being assessed, and the precision of the estimates of those populations.

## 2.1 Development of the mathematics assessments

PISA is a major source for internationally comparative results on the mathematics literacy (among other subjects) of students nearing the end of their legally required education.[4] It is conducted under the auspices of the Organization for Economic Cooperation and Development (OECD), and the PISA assessment framework is developed collaboratively with international mathematics experts and representatives from participating countries. The framework reflects developments in, and consensus on, the key content in mathematics that students need to

---

[3] In this report, the team of two key AIR research staff (Kim Gattis and Young Yee Kim) and two external experts (Mary Lindquist and Linda Dager Hall) is noted holistically as the "expert panel" and individually as the "panelists." The report refers to the external experts as "experts" and AIR staff (including additional support staff) as "research staff."

[4] The age when legally required, or compulsory, education ends ranges from 14 to 16 across the countries participating in PISA.

master to succeed in work or further education and that may be acquired inside or outside of school contexts.

NAEP is the main source of information on mathematics achievement (along with other subjects) in the United States at grades 4, 8, and 12, using nationally established benchmarks of performance (with ascribed achievement levels of *Basic*, *Proficient*, and *Advanced*). The National Assessment Governing Board establishes the mathematics framework and achievement levels based on the collaborative input of a wide range of experts and participants from government, education, business, and public sectors in the United States. The framework is intended to reflect the best thinking from these experts about the knowledge and skills needed by U.S. students in mathematics at these grade levels.

Thus, while PISA is deliberately constructed to reflect the mathematics content deemed important internationally throughout the collaborating countries, NAEP tailors its content to national practices acknowledged as important by experts and educators in the United States.

Both PISA and NAEP are conducted regularly to allow the monitoring of student outcomes over time. PISA uses a 3-year cycle in which, among its three key subjects (reading, mathematics, and science literacy), there is one major domain and two minor domains in each cycle. Thus, mathematics is assessed every 3 years, but as a major domain every 9 years, including most recently in 2012. The NAEP mathematics assessment is conducted every 2 years at grades 4 and 8 and about every 4 years at grade 12.[5] NAEP does not make any "major/minor domain" distinctions in its mathematics assessments, and when a grade is assessed, the full assessment is always given.

## 2.2 Underlying approach to assessing mathematics

Both the PISA and NAEP assessments are developed according to detailed test specifications that define their approaches to assessing mathematics and describe how the assessments are to be organized. However, despite some apparent similarities in framework organization—for example, both specify a range of content knowledge and cognitive skills, as shown in exhibit 2-1—PISA and NAEP differ in their underlying approaches to the assessment of mathematics. PISA's focus is on mathematics literacy—specifically, the application of mathematics in a wide variety of contexts, not limited to what is learned or applied in school—while NAEP's focus is more closely aligned with school-based curricular attainment that is sometimes described as a "union of curricula" across states.

[5] Mathematics assessments began in 1990 for all three grades. Prior to 2003, assessments for grades 4 and 8 were less frequent than every 2 years. Grade 12 mathematics assessments were administered in 1990, 1992, 1996, 2000, 2005, 2009, and 2013.

**Exhibit 2-1. Overview of PISA 2012 and NAEP 2013 mathematics frameworks, by framework dimension**

| Framework dimensions | PISA 2012 | NAEP 2013 (Grades 8 and 12) |
|---|---|---|
| Content | Content knowledge categories and description | Content areas and subtopics |
| | • Quantity<br>Quantifying attributes of objects, relationships, situations, and entities in the world; understanding various representations of those quantifications; and judging interpretations and arguments based on quantity | • Number properties and operations<br>– Number sense<br>– Estimation<br>– Number operations<br>– Ratios and proportional reasoning<br>– Properties of number and operations<br>– Mathematical reasoning using number |
| | • Space and shape<br>Understanding perspective (for example, in paintings), creating and reading maps, transforming shapes with and without technology, interpreting views of three-dimensional scenes from various perspectives, and constructing representations of shapes | • Measurement<br>– Measuring physical attributes<br>– Systems of measurement<br>– Measurement in triangles<br>• Geometry<br>– Dimension and shape<br>– Transformation of shapes and preservation of properties<br>– Relationships between geometric figures<br>– Position, direction, and coordinate geometry<br>– Mathematical reasoning in geometry |
| | • Uncertainty and data<br>Recognizing the place of variation in processes, having a sense of the quantification of that variation, acknowledging uncertainty and error in measurement, and knowing about chance | • Data analysis, statistics, and probability<br>– Data representation<br>– Features of datasets<br>– Experiments and samples<br>– Probability<br>– Mathematical reasoning with data |
| | • Change and relationships<br>Modeling change and relationships with appropriate functions and equations, as well as creating, interpreting, and translating among symbolic and graphical representations of relationships | • Algebra<br>– Patterns, relations, and functions<br>– Algebraic representations<br>– Variables, expressions, and operations<br>– Equations and inequalities<br>– Mathematical reasoning in algebra |

**Exhibit 2-1. Overview of PISA 2012 and NAEP 2013 mathematics frameworks, by framework dimension—*Continued***

| Framework dimension | PISA 2012 | NAEP 2013 (Grades 8 and 12) |
|---|---|---|
| Cognitive | No corresponding dimension | Mathematical complexity<br>• Low<br>• Moderate<br>• High |
| | Mathematical processes<br>• *Formulating* situations mathematically<br>• *Employing* mathematical concepts, facts, procedures, and reasoning<br>• *Interpreting*, applying, and evaluating mathematical outcomes | No corresponding dimension |
| | Mathematical capabilities<br>• Communication<br>• Mathematizing<br>• Representation<br>• Reasoning and argument<br>• Devising strategies for solving problems<br>• Using symbolic, formal, and technical language and operations<br>• Using mathematical tools | No corresponding dimension |
| Context | Context categories<br>• Personal<br>• Occupational<br>• Societal<br>• Scientific | No corresponding dimension |

*Note*: Geometry and measurement are combined at grade 12.
*Source*: NAEP mathematics framework (National Assessment Governing Board 2012) and PISA mathematics framework (OECD 2013).

The purpose of PISA is to measure how well education systems prepare students to apply their learned knowledge and skills in the context of real-world situations. This is done by assessing students' literacy in mathematics (as well as in reading and science) near the end of compulsory schooling. PISA defines mathematics literacy as "the capacity to formulate, employ, and interpret mathematics in a variety of contexts" (OECD 2013). It includes reasoning and using mathematical tools, concepts, facts, and procedures to describe, explain, and predict phenomena. For example, PISA would include items that assess how well students can take a real-world situation, such as a revolving door, and then use mathematics to describe the angles formed by the doors. As a result of this approach, the PISA assessment would not likely include any mathematics that is not set in a real-world context.

The purpose of NAEP is to measure academic achievement. In the NAEP mathematics framework, this is expressed as "students' knowledge of mathematics and the ability to apply that knowledge in problem-solving situations" (National Assessment Governing Board 2012). This approach, which emphasizes both students' knowledge and their ability to apply it, might be characterized as "school mathematics," or the mathematics that is typically taught in U.S. schools. For example, NAEP would include items that assess how well students can perform computations or simplify algebraic expressions, as well as the use of those skills when solving a word problem. There is an emphasis on mathematical skills—in some cases, devoid of any real-world context—as well as on the application of those skills to solve problems that are set in real-world contexts. This approach is similar to the way mathematics is typically addressed in U.S. schools, with sets of concepts and skills often introduced outside of a context and then those ideas brought to bear to solve contextualized problems.

The results of the in-depth review of item features, described in section 5 of this report, explore the differences stemming from these fundamentally different approaches to the assessment of mathematics, as well as the areas of intersection.

## 2.3 Target populations being assessed

PISA and NAEP are both sample-based assessments, meaning that each assessment is administered to a sample of students (rather than to all students) and the results are generalized to a larger population. However, each assessment defines the population to which it is generalizing, and therefore from which the sample is drawn, differently. One distinction between PISA and NAEP is that PISA uses an age-based sample, whereas NAEP uses grade-based samples. These choices stem from the purpose of each program: NAEP, to report on student achievement based on what students learn in school at specific grades; and PISA, to report on the mathematics knowledge and skills acquired by students cumulatively over their school and life experiences up to age 15.

The PISA target population is all 15-year-old students in at least seventh grade. This includes all students who were 15 years and 3 months to 16 years and 2 months at the beginning of the testing period and who were enrolled in a public or private school, regardless of full- or part-time status. The majority of respondents in the U.S. PISA 2012 sample were in 10th grade (71 percent), but some were in the 11th (17 percent), 9th (12 percent), or another grade (less than 1 percent).

The NAEP target population, on the other hand, is all students in the 4th, 8th, and 12th grades in public or private schools. Therefore, the PISA results are for students who are mostly in grades between those being tested for NAEP (8th and 12th grades), and slightly closer in grade proximity to those taking the NAEP 8th-grade assessment because of the timing of the respective assessments (with PISA given earlier in the school year than NAEP). Because the target populations of PISA and NAEP do not align neatly, this study focuses on comparisons of PISA with both the NAEP grade 8 and grade 12 item pools.

## 2.4 Precision of estimates in each assessment

The precision of score estimates depends largely on the sample sizes drawn from the target populations. PISA and NAEP both aim to offer accurate and reliable measures of students' performance for major subgroups and provide trend information on students' performance over time. To this end, each assessment draws sufficient samples from the U.S. student population. Because NAEP also reports students' performance at the state level (and at the district level for those districts participating in the Trial Urban District Assessment), NAEP samples much larger numbers of U.S. students than does PISA.[6]

PISA 2012 included about 6,000 15-year-old students in the U.S. national sample (see table 2-1). The NAEP 2013 national sample consisted of individual state samples of public school students and a national sample of private school students. About 162,000 and 40,000 students participated in the NAEP 2013 mathematics assessment at grades 8 and 12, respectively. Consequently, NAEP measured U.S. students' performance with a higher level of precision (i.e., with smaller standard errors) than did PISA, allowing NAEP to detect smaller differences between subgroups.

**Table 2-1. Number of students and schools participating in PISA 2012 and NAEP 2013 at grades 8 and grade 12**

| Assessment | Number of students | Number of schools |
|---|---|---|
| PISA 2012 (15-year-olds) | 6,000 | 240 |
| NAEP 2013 (grade 8) | 162,000 | 7,200 |
| NAEP 2013 (grade 12) | 40,000 | 1,800 |

*Note:* Number of students were rounded to 1,000 and the number of schools were rounded to 100.
*Source:* U.S. Department of Education, National Center for Education Statistics, Program for International Student Assessment (PISA), 2012; U.S. Department of Education, National Center for Education Statistics, National Assessment of Education Progress (NAEP), 2013 Mathematics Assessment.

# 3. Framework Comparisons

The first component of the study compared the PISA 2012 and NAEP 2013 assessment frameworks in order to provide insight into the similarities and differences in how each program conceptualizes mathematics and the mathematics knowledge and skills each program intends to measure. The comparison included the PISA 2012 mathematics framework and the NAEP 2013 mathematics framework at grades 8 and 12 (i.e., the two NAEP grades that straddle the modal grade of PISA's 15-year-olds). The comparisons focused primarily on the descriptions of the *content* and *cognitive* dimensions, since these are the two dimensions common to both PISA and NAEP (see exhibit 2-1). The approach to examining context is described in section 4 of the report.

---

[6] State participation in NAEP is required at grades 4 and 8 and voluntary at grade 12. The Trial Urban District Assessment is limited to grades 4 and 8.

## 3.1 Methods

As an initial step, research staff reviewed the descriptions of the content dimensions of the PISA and NAEP frameworks and then paired each of PISA's main content categories with one or more content areas in the NAEP framework at grade 8 and grade 12, as appropriate.[7]

Next, at the expert panel meetings, the panelists independently reviewed the content descriptions of each of the frameworks and evaluated the similarity of the paired content areas using a 4-point rating scale (1 = substantially or wholly different; 2 = quite dissimilar, but with some overlap; 3 = quite similar, but with some differences; 4 = exactly or almost the same). Using the same 4-point scale, the expert panelists also independently reviewed and evaluated the similarity of PISA's and NAEP's cognitive dimensions and of the frameworks overall.

Following the independent review, the panelists discussed the ratings as a group. After the discussion, each of the panelists documented their final similarity ratings on a data collection form (see appendix B), making any changes they deemed appropriate. The final similarity ratings were assigned based on the consensus of at least three of the four panelists.

## 3.2 Results

Table 3-1 summarizes the final similarity ratings for the PISA 2012 framework and the NAEP 2013 framework at grade 8 and grade 12.[8] The table shows that there was not a great deal of overlap between PISA and NAEP for any of the dimensions or for the frameworks overall. The highest rating given was a 2, "quite dissimilar but with some overlap," to four of the five content area pairings and to the frameworks overall. The lowest rating given was to the pairing of PISA's space and shape with NAEP's geometry at grade 8 and grade 12, which were rated a 1, "substantially or wholly different." Similarly, the pairing of the cognitive dimensions—PISA's mathematical processes and NAEP's mathematical complexities—were given a rating of 1.

[7] The research staff who performed the content matching and preclassifications (described in section 4.1) were independent of the key research staff who participated in the panel. This allowed the latter to review the frameworks and items without predisposition from the planning phase.
[8] Throughout this study, grades 8 and 12 are treated separately. Although they share common overall content areas and complexity levels, each has a unique item pool based on specifications uniquely defined for that grade.

**Table 3-1. Final similarity ratings for the framework dimensions and content pairings for PISA 2012 and NAEP 2013 at grades 8 and 12**

| Framework dimension | PISA 2012 content categories | NAEP 2013 content areas | Similarity rating between PISA and NAEP: | |
|---|---|---|---|---|
| | | | Grade 8 | Grade 12 |
| Content | Quantity | Number properties and operations | 2 | 2 |
| | Space and shape | Geometry | 1 | 1 |
| | Space and shape | Measurement | 2 | 2 |
| | Uncertainty and data | Data analysis, statistics, and probability | 2 | 2 |
| | Change and relationships | Algebra | 2 | 2 |
| Cognitive | Mathematical processes | Mathematical complexity | 1 | 1 |
| Context | Four situational contexts | † | † | † |
| Overall | PISA framework as a whole | NAEP framework as a whole | 2 | 2 |

† Not applicable. NAEP does not have a context dimension in its framework; therefore, similarity ratings are not possible.

*Note*: Each NAEP content area was matched with one of four PISA content categories with likely similarities, and the expert panel then rated the similarity between the NAEP and PISA pairs based on a 4-point scale, with "1" meaning the pairs were substantially or wholly different; "2" meaning they were quite dissimilar but with some overlap; "3" meaning they were quite similar but with some differences; and "4" meaning they were exactly or almost the same. Although geometry and measurement are combined for scaling and reporting in grade 12 NAEP, there are separate specifications and items for each content area.

# 4. Comparison of Item Pools to Framework-Related Dimensions

The second component of the study was a comparison of the item pools of the PISA 2012 and NAEP 2013 grade 8 and grade 12 mathematics assessments on framework-related dimensions. Rather than a direct comparison, this component consisted of classifying each assessment's items in terms of their content, cognitive aspects, and context. In the case of the content and cognitive aspects, items from each assessment were classified to the other assessment's framework. For context, items from both assessments were classified using an independent rubric that focused on the extent to which students needed to interpret the context of each item to solve the problem.

Section 4 is organized as follows:

- Section 4.1 describes the methods used in the item pool comparisons.
- Sections 4.2 through 4.4 present the results in each of the three areas in which the item pools were compared (content, cognitive aspects, and context, respectively).

- Section 4.5 presents the results of a comparison of the item formats. Each framework contains item format specifications that are used to guide the development of the assessment and of the construction of the items. This comparison provides an important basis for the subsequent discussions of other item features in section 5.

## 4.1 Methods

The approach to the item classification tasks was generally similar regardless of direction (i.e., PISA to NAEP or NAEP to PISA), with some slight differences in implementation as described in the steps below.

1. ***Identify which items to classify***. Initially, the methodology called for all the mathematics items from both assessments to be classified into their counterpart's framework. It was possible to follow this methodology for the placement of the PISA items in the NAEP framework because of the high degree of specificity of the NAEP framework as well as NAEP's underlying approach to mathematics.[9] However, the expert panel judged that many NAEP items would not fit within PISA's underlying approach to mathematics and that it would be unnecessary to classify all of the NAEP items. Thus, the expert panel sorted the NAEP items into two categories: those with "potential fit" and those with "definitely no fit" to the PISA framework. The "potential fit" items were then subject to the classification exercise described next, whereas the "definitely no fit" items were excluded. This latter subset of items included grade 12 items that measured advanced high school mathematics (such as trigonometry) and items with no or little context.[10]
2. ***Preclassification of items into content categories***. Research staff preclassified all of the items selected in the first step—all of the PISA items and the subset of "potential fit" NAEP items—into the content areas or categories in the counterpart's framework, although at different levels of detail depending on the framework. For example, the NAEP items were preclassified into the PISA framework in the four major content categories, but not in PISA's other content dimension: content topics. The NAEP framework, on the other hand, provides detailed specifications at multiple levels: within content areas there are specified subtopics, and within subtopics there are specified objectives. The PISA items were preclassified into the NAEP framework at the greatest level of specificity possible. Decisions about the preclassification process were made based on what would be most useful in facilitating the next step of the process, given the resources available.[11]

---

[9] Both PISA's paper-and-pencil and computer-based items were included in the classifications. However, because the international and national reports scale and report these two types of items differently, this report follows suit. The results in the main body of the report focus on the paper-and-pencil items; appendix C provides background information and results for the computer-based items.
[10] The expert panel treated items with very minimal context (e.g., "John did this…") as not fitting the PISA framework.
[11] In the body of this report, data are generally aggregated to the major content area. Appendix D provides the results at the NAEP subtopic level.

3. ***Expert panel review and classifications into content, cognitive, and context dimensions***. Next, the panelists reviewed the preclassifications and item pools individually and made their own classifications. For the NAEP items, in addition to classifying each item into one of the four PISA content categories, the panelists also classified each item into one of the 15 PISA content topics.[12] The panelists also classified the items into cognitive categories (i.e., NAEP items into PISA's mathematical processes and PISA items into NAEP's mathematical complexity levels) and into a set of three categories developed to describe the role of context.
4. ***Overall fit rating for content***. In addition to placing the PISA and NAEP items in the counterpart's framework, the expert panel also assigned an overall fit rating for each item's content classification. The rating was on a 4-point scale—4 = fits perfectly or almost perfectly, 3 = fits quite well, 2 = fits a little, and 1 = does not fit at all—that was intended to provide a sense of the degree of match of the assessment's items to the counterpart's content categories. The panel did not provide fit ratings for the cognitive or context classifications, and the results of fit ratings for content were used mainly as a criterion for determining the final classification of items rather than as a set of stand-alone results.
5. ***Expert panel discussion and determination of final classifications and fit ratings***. As a final step, the expert panel discussed the individual classifications, fit ratings, and rationales for their initial decisions. When the four panelists were in agreement, there was little discussion. When there were differences, the discussion lasted longer, sometimes 5 to 15 minutes for each item. The purpose of the discussion was for the panel to develop a common understanding and interpretation of the items and framework categories, not to enforce a consensus. Throughout the discussion, each panelist independently made revisions or reclassifications as desired. The final classification decision was determined based on consensus—which was defined as agreement among at least three of the four panelists—and, for the content classifications, a rating of at least 3 on overall fit. When there was no consensus, an item was identified as "not matched."

## 4.2 Results of the content dimension comparisons

This section presents the results of the PISA-to-NAEP content comparisons first and the NAEP-to-PISA comparisons second (as does section 4.3 for the context comparisons). This is because, as the results will show, the PISA items tended to fit more easily to the NAEP framework than the NAEP items did to the PISA framework (particularly in content). It is also because the comparisons of the PISA items to the NAEP framework involved the full set of PISA items, whereas the comparisons to the PISA framework were based on a subset of the NAEP items. What this suggests is that what PISA measures in terms of content falls within the boundaries of what NAEP defines as important content, even if PISA items are more frequently contextualized than NAEP items. In contrast, many NAEP items were ruled out from being fit to the PISA framework because they measure advanced mathematics or are not contextualized in the way that PISA items are (and, of those NAEP items that were included in the classifications,

[12] Appendix E provides the results at the PISA topic level.

a relatively large percentage assessed content that was not explicitly identified in the PISA framework).

## *PISA to NAEP*

Overall, the PISA items matched easily to the NAEP framework in terms of content. All but three of PISA's 85 items (96 percent) were matched to one of the five NAEP content areas. This included all items from the change and relationship category and all but one each from the other three PISA content categories (see table 4-1). (The 85 items are the paper-and-pencil items; the results for the PISA computer-based items are in appendix C.)

An important consideration in the PISA-to-NAEP classifications is grade level, since the expert panel was comparing the PISA items to the NAEP framework at all three grades (4, 8, and 12), even though grade 4 assessment items were not compared with the PISA framework. Three-quarters of the PISA items (64 items, or 75 percent) matched to the framework at grade 8, 12 items (or 14 percent) to the framework at grade 12, and 6 items (or 7 percent) to the framework at grade 4 (see table 4-1).[13] Overall, the mathematics content measured by the PISA assessment is most comparable to the content of the NAEP grade 8 assessment.

Another important consideration is how the PISA items were distributed across the NAEP content categories. In other words, how well did the PISA content categories match with the NAEP content areas at the item level? The majority of PISA items in each content category matched to a single corresponding content area in the NAEP framework, although the percentages varied. The PISA quantity items showed the strongest correspondence with their potential counterpart: 77 percent of items matched to NAEP's number properties and operations content area, although the remaining items matched to two other content areas. Also as expected, most PISA uncertainty and data items (71 percent) matched to the data analysis, statistics, and probability content area in NAEP. PISA's space and shape items mapped exclusively to geometry (62 percent) or measurement (33 percent). Finally, items in the change and relationships content category showed the least correspondence with a single NAEP area, with 52 percent matched to algebra; 29 percent to data analysis, statistics, and probability; and 14 percent to number properties and operations.

---

[13] The expert panel classified six PISA items to the NAEP framework at grade 4 because they were perceived as best matching descriptions given at that level of the framework. While such classifications may seem surprising, it is possible because (1) there are some mathematics topics that begin in grade 4 and reappear in later grades with added depth and (2) panelists were directed to classify items at the lowest level of the framework that was appropriate.

**Table 4-1. Number (N) and percentage (%) distribution of PISA items across NAEP content areas, by PISA content category**

| PISA content category | Grade | Any NAEP content area | | Number properties and operations | | Geometry | | Measurement | | Data analysis, statistics, and probability | | Algebra | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | N | % | N | % | N | % | N | % | N | % | N | % |
| Change and relationships[1] | No match | 0 | 0 | † | † | † | † | † | † | † | † | † | † |
| | Match | 21 | 100 | 3 | 14 | 1 | 5 | 0 | 0 | 6 | 29 | 11 | 52 |
| | Grade 4 | 1 | 5 | 1 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Grade 8 | 16 | 76 | 2 | 10 | 1 | 5 | 0 | 0 | 6 | 29 | 7 | 33 |
| | Grade 12 | 4 | 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 19 |
| Space and shape[1] | No match | 1 | 5 | † | † | † | † | † | † | † | † | † | † |
| | Match | 20 | 95 | 0 | 0 | 13 | 62 | 7 | 33 | 0 | 0 | 0 | 0 |
| | Grade 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Grade 8 | 15 | 71 | 0 | 0 | 8 | 38 | 7 | 33 | 0 | 0 | 0 | 0 |
| | Grade 12 | 5 | 24 | 0 | 0 | 5 | 24 | 0 | 0 | 0 | 0 | 0 | 0 |
| Uncertainty and data[1] | No match | 1 | 5 | † | † | † | † | † | † | † | † | † | † |
| | Match | 20 | 95 | 4 | 19 | 0 | 0 | 1 | 5 | 15 | 71 | 0 | 0 |
| | Grade 4 | 2 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 10 | 0 | 0 |
| | Grade 8 | 15 | 71 | 4 | 19 | 0 | 0 | 1 | 5 | 10 | 48 | 0 | 0 |
| | Grade 12 | 3 | 14 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 14 | 0 | 0 |
| Quantity[2] | No match | 1 | 5 | † | † | † | † | † | † | † | † | † | † |
| | Match | 21 | 96 | 17 | 77 | 0 | 0 | 2 | 9 | 2 | 9 | 0 | 0 |
| | Grade 4 | 3 | 14 | 2 | 9 | 0 | 0 | 0 | 0 | 1 | 5 | 0 | 0 |
| | Grade 8 | 18 | 82 | 15 | 68 | 0 | 0 | 2 | 9 | 1 | 5 | 0 | 0 |
| | Grade 12 | 0 | 0 | 0 | *0* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

† Not applicable. PISA items that did not match to the NAEP framework cannot be associated with particular content areas in the NAEP framework.
[1] Total number of items in this category was 21.
[2] Total number of items in this category was 22.
*Note*: N = Number; % = Percentage. Detail may not sum to totals because of rounding.

These analyses can be expanded by looking at the breadth of coverage within the NAEP content areas: that is, to what extent were the subtopics in the individual NAEP content areas covered by the PISA items? Within each content area, PISA items did in fact cover a certain portion of the subtopics (see appendix D), although they did not cover the full range possible. The subtopics that PISA did not cover are as follows:

- *Within Number properties and operations*: estimation; and mathematical reasoning using numbers;
- *Within Geometry*: position, direction, and coordinate geometry; and mathematical reasoning in geometry;
- *Within Measurement*: measurement in triangles;
- *Within Data analysis, statistics, and probability*: experiments and samples; and mathematical reasoning with data;
- *Within Algebra*: mathematical reasoning in algebra.

These subtopics, which NAEP covers but PISA does not, may indicate an important difference in the mathematics content covered in the two assessments.

With respect to content, the final analysis is of the distribution of PISA items across NAEP's content areas to the distribution of NAEP items across these content areas. This analysis compares the PISA items matched to the NAEP framework at grade 8 with the NAEP grade 8 items and the PISA items matched to the NAEP framework at grade 12 with the NAEP grade 12 items.[14]

At grade 8, NAEP emphasized algebra over the other content areas: 30 percent of items were in algebra, with the remainder fairly evenly distributed across the other content areas (see table 4-2a). In contrast, only 11 percent of the PISA items matching to the grade 8 framework were identified as algebra items. Instead, the PISA items had a stronger emphasis on number properties and operations (33 percent) and data analysis, statistics, and probability (27 percent).

At grade 12, similar proportions of NAEP items and PISA items (that is, the 12 items that matched to the grade 12 framework) were concentrated in algebra (about one-third) and data analysis, statistics, and probability (one-quarter) (see table 4-2b). However, while the rest of the PISA items were concentrated in geometry (42 percent), with no items in number properties and operations or in measurement, the rest of the NAEP items were distributed across the three other content areas, with a larger concentration in geometry (19 percent) than in the other two areas (12 percent each). However, the results for grade 12 should be viewed with caution since the comparison is based on a very small subset of PISA items.

[14] This report does not examine in detail the six PISA items that were matched to the NAEP grade 4 framework.

**Table 4-2a. Number and percentage distribution of the 64 PISA items matched to the NAEP framework at grade 8 and the NAEP grade 8 items, by NAEP content area**

| NAEP content area | PISA | | NAEP | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Number properties and operations | 21 | 33 | 29 | 19 |
| Geometry | 9 | 14 | 26 | 17 |
| Measurement | 10 | 16 | 29 | 19 |
| Data analysis, statistics, and probability | 17 | 27 | 23 | 15 |
| Algebra | 7 | 11 | 46 | 30 |

*Note:* Detail may not sum to totals because of rounding.

**Table 4-2b. Number and percentage distribution of the 12 PISA items matched to the NAEP framework at grade 12 and the NAEP grade 12 items, by NAEP content area**

| NAEP content area | PISA | | NAEP | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Number properties and operations | 0 | 0 | 22 | 12 |
| Geometry | 5 | 42 | 37 | 19 |
| Measurement | 0 | 0 | 22 | 12 |
| Data analysis, statistics, and probability | 3 | 25 | 48 | 25 |
| Algebra | 4 | 33 | 62 | 32 |

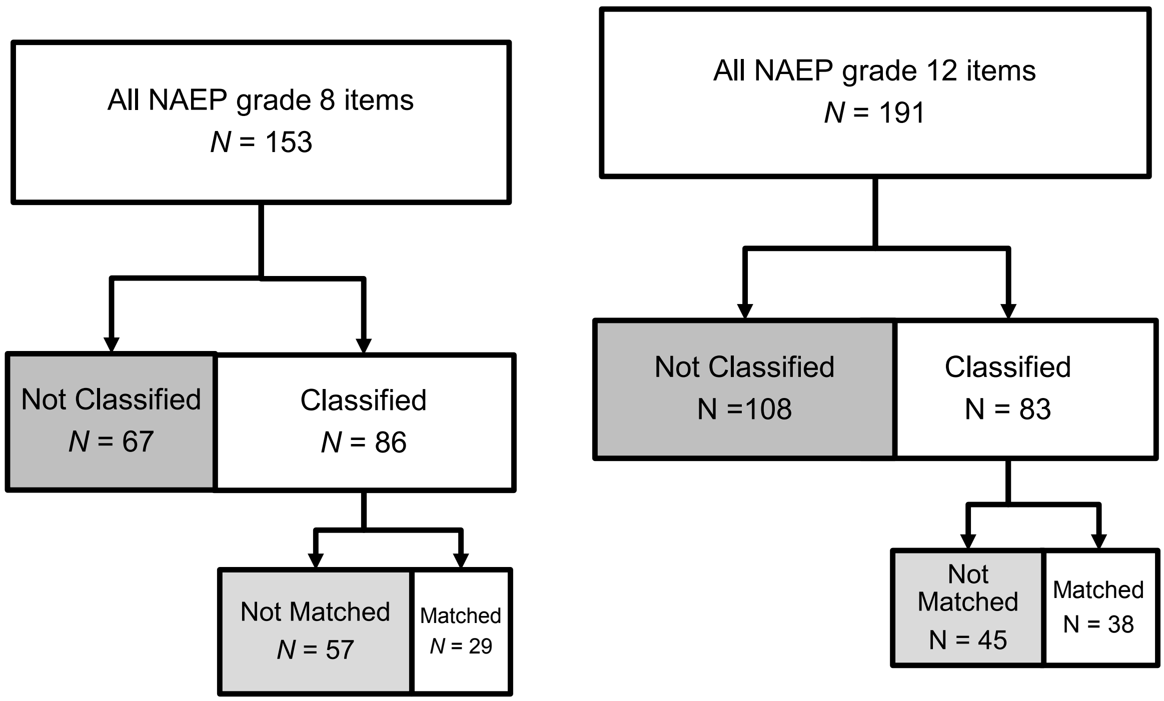
*Note*: Detail may not sum to totals because of rounding.

### *NAEP to PISA*

Whereas the PISA items matched easily to the NAEP framework, the opposite was not true. Not only were a large proportion of NAEP items excluded from the classification exercise as having "definitely no fit" to the PISA framework, substantial percentages of the items included in the classification exercise were determined not to match to the PISA framework.

Figure 4-1 provides a visual display of how the NAEP items were reduced from the full set through the study's two-step process: (1) determination of eligibility for the classification exercise (resulting in "classified" and "not classified" items); and (2) determination of a match to the PISA framework for the "classified" items (resulting in "matched" or "not matched"). (The same data, disaggregated by content area, are provided in tables 4-3a and 4-3b.) Most of the analyses in the NAEP-to-PISA sections of this report are based on either the subset of "classified" items or the subset of "matched" items.

**Figure 4-1. Determining eligibility of NAEP grade 8 and 12 items for classification and match to the PISA framework**




Eighty-six (or 56 percent) of NAEP's 153 grade 8 items and 83 (or 43 percent) of its 191 grade 12 items were identified as having a "potential fit" to the PISA framework and thus included in the expert panel's classification exercise (see tables 4-3a and 4-3b).[15] At both grades, items in data analysis, statistics, and probability and in measurement were most likely to qualify for the classification exercise, and geometry and algebra items were most likely to be excluded.

Of the 86 NAEP grade 8 items that were compared to the PISA framework, about one-third (29 items, or 34 percent) matched to one of the four PISA content categories, whereas about two-thirds (57 items, or 66 percent) did *not* match to any of PISA's content categories (table 4-3a). The matching rate varied by NAEP content area. Less than one-quarter of the classified NAEP grade 8 items in three of the five NAEP content areas matched to a PISA content category; number properties and operations had the lowest percentage of matching items (13 percent) and algebra the highest (70 percent).

---

[15] If calculated as a percentage of all NAEP items, instead of as a percentage of classified NAEP items, the matching rate for NAEP to PISA would be 19 percent (29 of 153) at grade 8 and 20 percent (38 of 191) for grade 12.

**Table 4-3a. Number and percentage of NAEP grade 8 items, by classification and match status to PISA**

| NAEP content area | Total number of items | Classification status | | | | Match status | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Not classified | | Classified | | Classified, but not matched to a PISA content category | | Classified and matched to a PISA content category | |
| | | Number | Percent | Number | Percent | Number | Percent | Number | Percent |
| All content areas | 153 | 67 | 44 | 86 | 56 | 57 | 66 | 29 | 34 |
| Number properties and operations | 29 | 14 | 48 | 15 | 52 | 13 | 87 | 2 | 13 |
| Geometry | 26 | 18 | 69 | 8 | 31 | 5 | 63 | 3 | 38 |
| Measurement | 29 | 8 | 28 | 21 | 72 | 16 | 76 | 5 | 24 |
| Data analysis, statistics, and probability | 23 | 1 | 4 | 22 | 96 | 17 | 77 | 5 | 23 |
| Algebra | 46 | 26 | 57 | 20 | 44 | 6 | 30 | 14 | 70 |

*Note:* Detail may not sum to totals because of rounding.

**Table 4-3b. Number and percentage of NAEP grade 12 items, by classification and match status to PISA**

| NAEP content area | Total number of items | Classification status | | | | Match status | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Not classified | | Classified | | Classified, but not matched to a PISA content category | | Classified and matched to a PISA content category | |
| | | Number | Percent | Number | Percent | Number | Percent | Number | Percent |
| All content areas | 191 | 108 | 57 | 83 | 43 | 45 | 54 | 38 | 46 |
| Number properties and operations | 22 | 14 | 64 | 8 | 36 | 8 | 100 | 0 | 0 |
| Geometry | 37 | 30 | 81 | 7 | 19 | 3 | 43 | 4 | 57 |
| Measurement | 22 | 10 | 46 | 12 | 55 | 6 | 50 | 6 | 50 |
| Data analysis, statistics, and probability | 48 | 7 | 15 | 41 | 85 | 25 | 61 | 16 | 39 |
| Algebra | 62 | 47 | 76 | 15 | 24 | 3 | 20 | 12 | 80 |

*Note:* Detail may not sum to totals because of rounding.

Of the 83 NAEP grade 12 items that were compared to the PISA framework, 38 items (or 46 percent) matched to a PISA content category, thus displaying a stronger correspondence than at grade 8 (see table 4-3b). The NAEP grade 12 items that matched least frequently to the PISA framework, again, were those in number properties and operations, none of which were matched to the PISA framework. The items that matched most frequently were those in the algebra content area (80 percent). From 39 to 57 percent of items in the other NAEP content areas were matched to the PISA framework.

Thus, at both grades 8 and 12, a large majority of the NAEP algebra items included in the classification exercise were matched to a PISA content category, while the matching rate was zero or very low for the number properties and operations items included in the classifications.

The next analysis examines how the matched NAEP items were distributed across the PISA content categories. In other words, how well did the NAEP content areas match the PISA content categories at the item level?

As shown in table 4-4a, at grade 8, a majority of items in each NAEP content area were matched to a single PISA content category. In three of the five content areas (number properties and operations; geometry; and data analysis, statistics, and probability), all of the NAEP items were matched to the corresponding PISA categories (quantity, space and shape, and uncertainty and data, respectively). The exceptions were the NAEP grade 8 measurement items, which were split between space and shape (60 percent) and quantity (40 percent), and the algebra items, which matched most closely to change and relationships (86 percent), but also had small percentages in the uncertainty and data and the quantity categories (7 percent each).

The NAEP grade 12 items that were matched to the PISA framework did not show quite as strong a correspondence between content areas at the item level as the grade 8 items did (see table 4-4b). Only NAEP's geometry items matched 100 percent to PISA's space and shape category. In three other content areas, the majority of the NAEP items were matched to the most closely corresponding PISA content category: 67 percent of NAEP measurement items to space and shape; 81 percent of data analysis, statistics, and probability items to uncertainty and data; and 92 percent of algebra items to change and relationships. Notably, only one-third of NAEP's measurement items matched with PISA's quantity category. None of the items in number properties and operations were matched to any PISA content category.

What is perhaps surprising about these results is that while the framework analysis showed the weakest overlap between NAEP's geometry content area and PISA's space and shape content category, the NAEP geometry items actually fit squarely within PISA's definition of space and shape. It is important, however, to note that less than one-third of NAEP's geometry items were included in the classification exercise (tables 4-3a and 4-3b). The results also show that what NAEP implements as a measurement item (at least among those items that were included in the classification exercise) can fall within PISA's definitions of either space and shape or quantity.

**Table 4-4a. Percentage distribution of the 29 NAEP grade 8 items matched to the PISA framework, by PISA content categories**

| NAEP content area | PISA content category | | | |
|---|---|---|---|---|
| | Change and relationships | Space and shape | Uncertainty and data | Quantity |
| Number properties and operations | 0 | 0 | 0 | 100 |
| Geometry | 0 | 100 | 0 | 0 |
| Measurement | 0 | 60 | 0 | 40 |
| Data analysis, statistics, and probability | 0 | 0 | 100 | 0 |
| Algebra | 86 | 0 | 7 | 7 |

*Note:* Detail may not sum to totals because of rounding.

**Table 4-4b. Percentage distribution of the 38 NAEP grade 12 items matched to the PISA framework, by PISA content categories**

| NAEP content area | PISA content category | | | |
|---|---|---|---|---|
| | Change and relationships | Space and shape | Uncertainty and data | Quantity |
| Number properties and operations | † | † | † | † |
| Geometry | 0 | 100 | 0 | 0 |
| Measurement | 0 | 67 | 0 | 33 |
| Data analysis, statistics, and probability | 13 | 0 | 81 | 6 |
| Algebra | 92 | 0 | 0 | 8 |

† Not applicable. No grade 12 items in the NAEP number properties and operations content area were matched to a PISA content category.
*Note:* Detail may not sum to totals because of rounding.

The final analysis with respect to content is a comparison of the distribution of NAEP items across PISA's content categories to the distribution of PISA items across the same categories. In other words, how do PISA and NAEP items compare in the relative emphasis placed on different content categories? Note, however, that this analysis is limited, because only the subset of NAEP items included in the classification exercise will have associated content category matches, while the full set of PISA items will by definition have matches. Therefore, this analysis shows how those NAEP items that were deemed most like PISA items compare to the PISA assessment in terms of content distribution.

The PISA items were fairly evenly distributed across the four categories (see table 4-5a), as the framework's guidelines suggest. In contrast, at grade 8, the percentage of NAEP items matching PISA's change and relationships category (14 percent) was twice as high as the percentage matching the other three categories (7 percent each in space and shape and in uncertainty and data and 6 percent in quantity).

At grade 12, the pattern varied slightly from that at grade 8, with equal percentages of NAEP items matching the change and relationships and the uncertainty and data categories (16 percent each) (see table 4-5b). The PISA content category with the lowest percentage of matching NAEP items was quantity (5 percent).

**Table 4-5a. Number and percentage distribution of PISA items and the 86 NAEP grade 8 items matched to the PISA framework, by PISA content category**

| PISA content category | PISA | | NAEP | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Change and relationships | 21 | 25 | 12 | 14 |
| Space and shape | 21 | 25 | 6 | 7 |
| Uncertainty and data | 21 | 25 | 6 | 7 |
| Quantity | 22 | 26 | 5 | 6 |
| Not matched | † | † | 57 | 66 |

† Not applicable.
*Note:* Detail may not sum to totals because of rounding.

**Table 4-5b. Number and percentage distribution of PISA items and the 83 NAEP grade 12 items matched to the PISA framework, by PISA content category**

| PISA content category | PISA | | NAEP | |
|---|---|---|---|---|
| | Number | Percentage | Number | Percentage |
| Change and relationships | 21 | 25 | 13 | 16 |
| Space and shape | 21 | 25 | 8 | 10 |
| Uncertainty and data | 21 | 25 | 13 | 16 |
| Quantity | 22 | 26 | 4 | 5 |
| Not matched | † | † | 45 | 54 |

† Not applicable.
*Note:* Detail may not sum to totals because of rounding.

## 4.3. Results of the cognitive dimension comparisons

As described in the Methods sections (sections 3.1 and 4.1), the PISA and NAEP items were also compared to the cognitive dimensions of the counterpart framework. Thus, the PISA items were categorized into NAEP's complexity levels (low, moderate, and high) and the NAEP items were categorized into PISA's mathematical processes (employ, formulate, and interpret).

### *PISA to NAEP*

All but one of PISA's 85 items could be categorized into one of NAEP's complexity levels,[16] although very few were categorized as high complexity. Across all categories, about half (45 items, or 53 percent) were categorized as moderate complexity, slightly fewer (37 items, or 44 percent) were categorized as low complexity, and just 2 items (2 percent) were categorized as high complexity (see table 4-6).

[16] One PISA item was not classified into any of NAEP's complexity levels due to a lack of a classification agreement among the panelists.

There were differences in the percentage distribution across complexity levels by content category. For example, PISA's quantity items were more likely to be categorized as low complexity (64 percent) than were items from the other categories (each less than 50 percent). At the other end of the distribution, the only PISA items categorized as high complexity were the uncertainty and data items. PISA's space and shape content category had the highest percentage of items categorized as moderate complexity (76 percent) and the lowest percentage categorized as low complexity (24 percent).

**Table 4-6. Number (N) and percentage (%) distribution of PISA items, by NAEP complexity level**

| PISA content category | Total number of items | NAEP complexity level | | | | | | Not matched | |
|---|---|---|---|---|---|---|---|---|---|
| | | Low | | Moderate | | High | | | |
| | | N | % | N | % | N | % | N | % |
| All categories | 85 | 37 | 44 | 45 | 53 | 2 | 2 | 1 | 1 |
| Change and relationships | 21 | 8 | 38 | 13 | 62 | 0 | 0 | 0 | 0 |
| Space and shape | 21 | 5 | 24 | 16 | 76 | 0 | 0 | 0 | 0 |
| Uncertainty and data | 21 | 10 | 48 | 8 | 38 | 2 | 10 | 1 | 5 |
| Quantity | 22 | 14 | 64 | 8 | 36 | 0 | 0 | 0 | 0 |

*Note:* N = Number; % = Percentage. Detail may not sum to totals because of rounding

The final analysis related to the cognitive dimensions compares the distribution of PISA items across NAEP mathematical complexity levels with the distribution of NAEP items across these levels.

Comparing the distribution of PISA and NAEP items across the NAEP mathematical complexity levels identifies a number of differences. Overall, NAEP had higher percentages of items (56 percent at both grades 8 and 12) categorized as low complexity than PISA (44 percent) did (see table 4-7a). Conversely, PISA had a higher percentage of items (53 percent) categorized as moderate complexity than NAEP (39 percent at both grades) did. In both assessments, the percentage of items categorized as high complexity was low: 2 percent of PISA items and 5 percent of both NAEP grade 8 and grade 12 items. Overall, the PISA item pool seems to be of generally higher complexity than the NAEP item pool.

The initial process of reducing the NAEP items to a subset for the classification exercise was based mainly on one aspect of the items, namely, whether there is context or not (with those having no context—or, "naked" items—excluded from the classifications). As these items were generally more straightforward in terms of tasks, it is not surprising that they had lower mathematical complexity than the remaining items reported on here.

Focusing, instead, on the subset of NAEP items included in the classification exercise (i.e., those initially considered to have a "potential fit" with PISA's overall approach to mathematics), the picture is quite different. Among these items, at both grades, NAEP had smaller percentages of items categorized as low complexity and higher percentages categorized as moderate or high complexity than PISA and the full set of NAEP items. That is, the subset of NAEP items deemed most likely to fit the PISA framework was generally of higher complexity than the PISA item pool and the overall NAEP item pool as well.

At grade 8, about one-quarter of the 86 classified NAEP items were categorized as low complexity and about two-thirds were classified as moderate complexity. The percentage of the classified NAEP items in the high complexity category was 10 percent, compared to 5 percent for the full item set (see tables 4-7a and 4-7b). At grade 12, the pattern was similar. Among the 83 classified items, NAEP had the largest percentage of items in the moderate complexity level (58 percent), followed by the low complexity level (34 percent) and high complexity level (8 percent) (see table 4-7c).

**Table 4-7a. Percentage distribution of PISA and NAEP grade 8 and grade 12 items, by NAEP complexity level**

| Item Pool | NAEP complexity level | | | |
|---|---|---|---|---|
| | Low | Moderate | High | Not matched |
| PISA | 44 | 53 | 2 | 1 |
| NAEP grade 8 | 56 | 39 | 5 | † |
| NAEP grade 12 | 56 | 39 | 5 | † |

† Not applicable
*Note:* Detail may not sum to totals because of rounding.

**Table 4-7b. Percentage distribution of PISA items and the 29 classified NAEP grade 8 items across NAEP mathematical complexity levels, by PISA content category**

| PISA content category | NAEP complexity level | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Low | | Moderate | | High | | Not matched | |
| | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) |
| Any content category | 44 | 24 | 53 | 66 | 2 | 10 | 1 | † |
| Change and relationships | 38 | 42 | 62 | 42 | 0 | 17 | 0 | † |
| Space and shape | 24 | 33 | 76 | 67 | 0 | 0 | 0 | † |
| Uncertainty and data | 48 | 0 | 38 | 83 | 10 | 17 | 5 | † |
| Quantity | 64 | 0 | 36 | 100 | 0 | 0 | 0 | † |

† Not applicable
*Note:* Detail may not sum to totals because of rounding.

**Table 4-7c. Percentage distribution of PISA items and the 38 classified NAEP grade 12 items across NAEP mathematical complexity levels, by PISA content category**

| PISA content category | NAEP complexity level | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Low | | Moderate | | High | | Not matched | |
| | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) |
| Any content category | 44 | 34 | 53 | 58 | 2 | 8 | 1 | † |
| Change and relationships | 38 | 54 | 62 | 46 | 0 | 0 | 0 | † |
| Space and shape | 24 | 0 | 76 | 75 | 0 | 25 | 0 | † |
| Uncertainty and data | 48 | 46 | 38 | 54 | 10 | 0 | 5 | † |
| Quantity | 64 | 0 | 36 | 75 | 0 | 25 | 0 | † |

† Not applicable
*Note:* Detail may not sum to totals because of rounding.

## *NAEP to PISA*

The classified NAEP items showed a much stronger correspondence with the PISA framework's cognitive dimension than they did with its content dimension. All of the NAEP items in the classification exercise, except for one grade 8 item,[17] could be placed within one of PISA's three mathematical processes: employ, formulate, and interpret (see tables 4-8a and 4-8b).

At grade 8, about two-thirds of the classified NAEP items (57 items, or 66 percent) were matched to the "employ" category, with about one-quarter (20 items, or 23 percent) matched to the "formulate" category and one-tenth (8 items, or 9 percent) to the "interpret" category (see table 4-8a). However, there were significant differences by content area. For example, all of the NAEP grade 8 number properties and operations items were matched to the "employ" category, compared to less than one-third of the algebra items (30 percent). Exactly half of the algebra

[17] One NAEP item was not classified to any of PISA's three mathematical processes due to a lack of a classification agreement among the panelists.

items were matched to the “formulate” category. The geometry items were matched evenly between the “employ” and “formulate” categories (50 percent in each), with no items in the “interpret” category. The NAEP data analysis, statistics, and probability content area had the highest percentage of items in the “interpret” category (18 percent).

At grade 12, the overall item distribution was similar to that at grade 8. Close to two-thirds of the NAEP grade 12 items in the classification exercise (51 items, or 61 percent) matched to the “employ” category, with about one-quarter (22 items, or 27 percent) matching to the “formulate” category and one-tenth (10 items, or 12 percent) to the “interpret” category (see table 4-8b). There were variations by content area similar to those at grade 8. As at grade 8, all of the grade 12 number properties and operations items matched to the “employ” category and, for each content area except algebra, the majority of items matched here as well: 83 percent of measurement, 57 percent of geometry, and 56 percent of data analysis, statistics, and probability items. The majority of algebra items (60 percent) matched to the “formulate” category. At grade 12, the geometry content area had the highest percentage of items in the “interpret” category (29 percent); in contrast, at grade 8, no geometry items were in this category.

**Table 4-8a. Number (N) and percentage (%) of the 86 NAEP grade 8 items classified to the PISA framework, by PISA mathematical process**

| NAEP content area | Total number of items | PISA mathematical process | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Employ | | Formulate | | Interpret | | Not matched | |
| | | N | % | N | % | N | % | N | % |
| All content areas | 86 | 57 | 66 | 20 | 23 | 8 | 9 | 1 | 1 |
| Number properties and operations | 15 | 15 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| Geometry | 8 | 4 | 50 | 4 | 50 | 0 | 0 | 0 | 0 |
| Measurement | 21 | 16 | 76 | 4 | 19 | 1 | 5 | 0 | 0 |
| Data analysis, statistics, and probability | 22 | 16 | 73 | 2 | 9 | 4 | 18 | 0 | 0 |
| Algebra | 20 | 6 | 30 | 10 | 50 | 3 | 15 | 1 | 5 |

Note: N = Number; % = Percentage. Detail may not sum to totals because of rounding.

**Table 4-8b. Number (N) and percentage (%) of the 83 NAEP grade 12 items classified to the PISA framework across PISA mathematical processes, by NAEP content area**

| NAEP content area | Total number of items | PISA mathematical process | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Employ | | Formulate | | Interpret | | Not matched | |
| | | N | % | N | % | N | % | N | % |
| All content areas | 83 | 51 | 61 | 22 | 27 | 10 | 12 | 0 | 0 |
| Number properties and operations | 8 | 8 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| Geometry | 7 | 4 | 57 | 1 | 14 | 2 | 29 | 0 | 0 |
| Measurement | 12 | 10 | 83 | 2 | 17 | 0 | 0 | 0 | 0 |
| Data analysis, statistics, and probability | 41 | 23 | 56 | 10 | 24 | 8 | 20 | 0 | 0 |
| Algebra | 15 | 6 | 40 | 9 | 60 | 0 | 0 | 0 | 0 |

*Note:* N = Number; % = Percentage. Detail may not sum to totals because of rounding.

The final analysis, in terms of the cognitive dimensions, compares the distribution of NAEP items across PISA's mathematical processes with the distribution of PISA items across those processes. Again, for NAEP, this is based on just those items that were included in the classification exercise.

The item distributions differed between PISA and NAEP at both grades in a similar pattern (see tables 4-9a and 4-9b). Namely, the NAEP items classified to the PISA framework matched more frequently to the "employ" category (66 percent at grade 8 and 61 percent at grade 12) than the PISA items (44 percent). Conversely, the NAEP items classified to the PISA framework matched less frequently to the "interpret" category (9 percent at grade 8 and 12 percent at grade 12) than the PISA items (25 percent). It is not known how the distributions of NAEP items would change if all the items could have been included in the classification exercise.

**Table 4-9a. Number and percentage distribution of PISA items and the 86 NAEP grade 8 items classified to the PISA framework, by PISA mathematical process**

| PISA mathematical process | PISA | | NAEP | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Employ | 37 | 44 | 57 | 66 |
| Formulate | 27 | 32 | 20 | 23 |
| Interpret | 21 | 25 | 8 | 9 |
| Not matched | 0 | 0 | 1 | 1 |

*Note:* Detail may not sum to totals because of rounding.

**Table 4-9b. Number and percentage distribution of PISA items and the 83 NAEP grade 12 items classified to the PISA framework, by PISA mathematical process**

| PISA mathematical process | PISA | | NAEP | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Employ | 37 | 44 | 51 | 61 |
| Formulate | 27 | 32 | 22 | 27 |
| Interpret | 21 | 25 | 10 | 12 |
| Not matched | 0 | 0 | 0 | 0 |

*Note:* Detail may not sum to totals because of rounding.

## 4.4 Comparing contexts

As described previously, the expert panel also examined the item pools in terms of context. Context, as described in assessment frameworks, has generally been limited to describing the settings in which items are situated. For example, PISA has four context categories that span students' experiences, beginning with the personal (e.g., shopping, travel, or personal health) and expanding to the occupational (e.g., measuring, accounting, or design), societal (e.g., public transport or national statistics), and scientific (e.g., medicine, weather, or mathematics itself). It is not surprising, then, that nearly all of the PISA items are framed in an identifiable situational context. While the NAEP framework does not specify particular contexts of interest, when items are placed in a context, it can be either a mathematical context or a real-world application.

Context, however, can also play a more fundamental role in an assessment item than providing situational breadth if it needs to be taken into account when a student responds to the item. To describe the degree to which context plays such a role in the functioning of an item, two mathematics experts compared items from each assessment to a three-part rubric that was developed explicitly for this purpose (see figure 4-2). At one end are purely mathematical items that have "no context." In the middle are those items that demand very little in the way of mathematizing,[18] but nonetheless are in an identifiable real-word context ("contextualized"). At the other end are those items that demand more interpretation of the situation ("context dependent"). The final decision on the context category for each item was made based on a consensus of the experts. To illustrate these definitions, appendix F provides examples of PISA and NAEP items classified in each of the context categories. All PISA 2012 items were classified to this rubric, as were the 86 NAEP grade 8 and 83 NAEP grade 12 items included in the classification exercise.

[18] For a definition and discussion of "mathematizing," see section 5.1.

**Figure 4-2. Context classification rubric**

| Context category | Description |
|---|---|
| No context | These items have not been placed in any context other than a purely mathematical one. |
| Contextualized | These items have been placed in a context. However, the context is not strongly relevant to finding the correct answer. The context could be easily interchanged with another one. Moreover, the amount of information that students must extract from the context to solve the problem is low, and little interpretation of the context is required to find the answer. For items that present problems to solve, the translation from context to mathematical model is direct; students do not need to make assumptions or decisions related to the context. For items that require students to explain their answer, an item classified in this category would require an explanation that is mathematical and without reference to the context. |
| Context dependent | These items have been placed in a context that affects the way students think about the problems. Items in this category require students to continually think about how to interpret the context and possibly to use that context to interpret results. Students may need to make some assumptions to answer the question. Solving these items requires more interaction, more back-and-forth, between the mathematics and the context. |

The primary distinction between PISA and NAEP mathematics items, as suggested earlier, is that only one of the PISA items was categorized as having no context, whereas 6 percent of NAEP's grade 8 classified items and 11 percent of its grade 12 classified items were categorized in this way (see table 4-10). About two-thirds of PISA items, rather, are categorized as contextualized, and about one-third are classified as context dependent.

NAEP has larger percentages of items at both grade 8 and grade 12 that are categorized as contextualized than does PISA, but smaller percentages of items at both grades categorized as context dependent. Grade 8 has a higher percentage of contextualized items (80 percent) than does grade 12 (72 percent). It must be kept in mind, however, that these comparisons are based on only those NAEP items that were included in the classification exercise. If the excluded items were to be included in these analyses—given that the exclusions were largely for a lack of context—there would likely be a much larger percentage of NAEP items classified as having no context and smaller percentages in the other two categories.

**Table 4-10. Number and percentage distribution of the PISA 2012 items and the classified NAEP 2013 items at grade 8 and grade 12, by context category**

| | PISA | | NAEP grade 8 | | NAEP grade 12 | |
|---|---|---|---|---|---|---|
| **Context** | **Number** | **Percent** | **Number** | **Percent** | **Number** | **Percent** |
| Total | 85 | 100 | 86 | 100 | 83 | 100 |
| No context | 1 | 1 | 5 | 6 | 9 | 11 |
| Contextualized | 58 | 68 | 69 | 80 | 60 | 72 |
| Context dependent | 26 | 31 | 12 | 14 | 14 | 17 |

*Note:* Detail may not sum to totals because of rounding.

Looking at the results by content area shows that there was less variation among the PISA items than among the NAEP items (see tables 4-11, 4-12a, and 4-12b). The only PISA item categorized as having no context was a space and shape item. Otherwise, about two-thirds of the space and shape items, as well as the uncertainty and data items, were contextualized and one-third were context dependent, similar to the distribution of items overall (table 4-10). The slight differences were among the change and relationship items and the quantity items, with the former more likely to be context dependent than items in any other content category and the latter more likely to be contextualized than items in any other content category.

**Table 4-11. Number and percentage distribution of PISA 2012 items, by PISA content category and context category**

| Context | Change and relationships | | Space and shape | | Uncertainty and data | | Quantity | |
|---|---|---|---|---|---|---|---|---|
| | Number | Percent | Number | Percent | Number | Percent | Number | Percent |
| Total | 21 | 100 | 21 | 100 | 21 | 100 | 22 | 100 |
| No context | 0 | 0 | 1 | 5 | 0 | 0 | 0 | 0 |
| Contextualized | 13 | 62 | 14 | 67 | 14 | 67 | 17 | 77 |
| Context dependent | 8 | 38 | 6 | 29 | 7 | 33 | 5 | 23 |

*Note:* Detail may not sum to totals because of rounding.

Among the NAEP items included in the classification exercise, there appears to more variation in context by content area than in PISA (see tables 4-12a and 4-12b). At both grade 8 and grade 12, the NAEP geometry items were the most likely to be categorized as having no context, followed by the measurement items. In contrast, relatively large percentages of data analysis, statistics, and probability items at both grades (18 and 20 percent, respectively) were categorized as context dependent, compared to the items in most other areas at both grades. No number properties and operations items at either grade were considered context dependent. There also were relatively large percentages of grade 8 geometry items (25 percent) and grade 12 algebra items (20 percent) categorized as context dependent.

**Table 4-12a. Number (N) and percentage (%) distribution of the 86 classified NAEP 2013 grade 8 items, by content area and context category**

| Context | Number properties and operations | | Geometry | | Measurement | | Data analysis, statistics, and probability | | Algebra | |
|---|---|---|---|---|---|---|---|---|---|---|
| | N | % | N | % | N | % | N | % | N | % |
| All items | 15 | 100 | 8 | 100 | 21 | 100 | 22 | 100 | 20 | 100 |
| No context | 0 | 0 | 2 | 25 | 3 | 14 | 0 | 0 | 0 | 0 |
| Contextualized | 15 | 100 | 4 | 50 | 15 | 71 | 18 | 82 | 17 | 85 |
| Context dependent | 0 | 0 | 2 | 25 | 3 | 14 | 4 | 18 | 3 | 15 |

*Note:* N = Number; % = Percentage. Detail may not sum to totals because of rounding.

**Table 4-12b. Number (N) and percentage (%) distribution of the 83 classified NAEP 2013 grade 12 items, by content area and context category**

| Context | Number properties and operations | | Geometry | | Measurement | | Data analysis, statistics, and probability | | Algebra | |
|---|---|---|---|---|---|---|---|---|---|---|
| | N | % | N | % | N | % | N | % | N | % |
| All items | 8 | 100 | 7 | 100 | 12 | 100 | 41 | 100 | 15 | 100 |
| No context | 0 | 0 | 4 | 57 | 3 | 25 | 2 | 5 | 0 | 0 |
| Contextualized | 8 | 100 | 2 | 29 | 7 | 58 | 31 | 76 | 12 | 80 |
| Context dependent | 0 | 0 | 1 | 14 | 2 | 17 | 8 | 20 | 3 | 20 |

*Note:* N = Number; % = Percentage. Detail may not sum to totals because of rounding.

## 4.5 Comparing item formats

The analysis for this section was conducted based on item classification data provided by test developers for each assessment. Both PISA and NAEP use three item formats in their mathematics assessments, and although the names of these formats are slightly different, they are similarly defined (see table 4-13). The PISA framework specifies that approximately equal numbers of items be used across the three item formats in constructing an assessment. The NAEP framework specifies that multiple-choice items and constructed-response items (both short and extended) have about the same amount of testing time.

As shown in table 4-13, the distribution of PISA 2012 items across the three specified formats is fairly balanced, as the framework suggests. In contrast, more than two-thirds of NAEP 2013 items are multiple-choice items.[19] However, as section 5 shows, the more substantial differences are in approaches to scoring the constructed-response items.

[19] Although this appears to deviate from the framework targets, the framework targets for NAEP are given in percentages of testing time, whereas the comparisons in this report are based on percentages of items.

**Table 4-13. Item format descriptions and percentage distribution of PISA 2012 and NAEP 2013 (grade 8 and grade 12) items, by format**

| Item format (PISA name /NAEP name) | Description | Percentage of items | | |
|---|---|---|---|---|
| | | PISA | NAEP grade 8 | NAEP grade 12 |
| Selected response/ Multiple choice | Requires student to select from a number of response options | 39 | 75 | 69 |
| Closed constructed response/Short constructed response | Requires student to generate a response (e.g., provide a number or a category) that can sometimes be scored automatically or by using a simple scoring rubric | 34 | 22 | 25 |
| Open constructed response/Extended constructed response | Requires student to generate a longer response, often requiring an explanation or justification, that must be scored manually | 27 | 3 | 7 |

*Note:* Detail may not sum to totals because of rounding.

# 5. In-Depth Review of Item Features

The third component of the study was an in-depth item review to compare the features of the PISA 2012 and NAEP 2013 grade 8 and grade 12 items. This is a new aspect of NCES-sponsored assessment comparison studies; it was added to provide a more in-depth understanding of the *way* in which the different assessment items measure mathematics, not just *what* they measure in terms of knowledge and skills. Essentially, it is meant to explore the question of whether items from different assessments that have been classified similarly in content are actually measuring that content (or process or complexity) in the same way or whether there are important differences based on the items’ features.

An initial list of potentially important item features was developed by research staff and then revised based on the expert panel discussions throughout the item pool classification process. The panelists then determined six item features (see figure 5-1) that they felt would increase understanding of how the NAEP and PISA item pools differ beyond the differences of the major framework classifications. The panel then selected clusters of released NAEP and PISA items to illustrate how items classified to the same framework dimension may not necessarily measure that dimension in the same way. This section is organized by the observations related to the six selected features, with examples provided when appropriate.

**Figure 5-1. Overview of the features considered in the in-depth review of PISA 2012 and NAEP 2013 items**

| Item features | Description |
|---|---|
| Mathematizing in real-world problem solving | Mathematizing is the degree to which students must translate a real-world problem situation into a mathematical problem. |
| Role of context | Here, context relates to the context in which a problem is situated—not in terms of what it is, but in terms of the extent to which it must be accounted for in solving the problem. |
| Reading load | Reading load refers to the amount of reading required by the stimulus material and item. |
| Use of representations | Representations are those visuals that are not strictly text (e.g., tables, graphs, illustrations) that are part of the stimulus material for an item or items. |
| Clustering and scaffolding | Clustering refers to the grouping of items with common stimulus material, and scaffolding refers to the building of those items, and potentially student' responses, on one another sequentially. |
| Scoring of constructed-response items | Here, scoring refers to the manner in which students earn credit on the items in which they are required to provide a response rather than select a response. |

As described in the overview of the assessments (see section 2), there is a fundamental difference between the PISA and NAEP assessments in their approach to mathematics. PISA is based on mathematics literacy, NAEP on school mathematics. However, as this section will show, these two approaches are not mutually exclusive. Certain types of NAEP's word problems are intended to depict real-world situations, so NAEP and PISA will show some overlap. But the fundamental aim of the two assessments does differ, with NAEP including mathematical content that might be context-free and PISA using contexts as the starting point for assessment. This difference plays out in the way the frameworks are organized and in the mathematical activities that are assessed, as will be seen below.

## 5.1 Mathematizing in real-world problem solving

As noted above, PISA emphasizes mathematics as it is encountered in the context of real-world situations, which includes solving real-world problems. NAEP also measures students' abilities to apply mathematical knowledge and skills to solve real-world problems, but also assesses the acquisition of those skills, or the understanding of a concept, without a context or within a mathematical context. Considering real-world problem solving as a subset of each assessment, the panel identified another important difference about what these assessments are measuring: the degree of *mathematizing* required.

*Mathematizing* is essentially the work that applied mathematicians do. It is the process of making mathematical sense of a real-world situation. Mathematizing comes into play when mathematical tools or concepts are brought to bear to describe, explain, or predict relationships that occur in the world of experience. While there are similarities between solving real-world problems and mathematizing, there are also some differences. Consider the following example from PISA:

*The Town Council has decided to construct a streetlight in a small triangular park so that it illuminates the whole park. Where should the light be placed?*

*Source: PISA 2003 assessment framework.*

The example begins with a problem that is situated in reality. In the process of mathematizing, the first step is to organize the problem according to mathematical concepts (represent the park with a triangle, the streetlight as the center of a circle of illumination). Gradually trim away "reality" through processes, such as assumptions about essential features of the problem, until its mathematical essence is revealed (the problem can be translated to locating the center of a circle that circumscribes the triangle). Now solve the mathematical problem (e.g., the center is at the intersection of the perpendicular bisectors of the sides of the triangle). Finally, make sense of the solution in terms of the reality of the problem (i.e., the solution works if all the angles of the triangle are acute, but otherwise the location would be outside the park, or consider that there may be a tree or building in the desired location for the streetlight).

Contrast this with the usual steps of problem solving: understand the problem, make a plan, do the plan, and check the solution (Polya, 1945). This heuristic was intended to be useful to teach students how to solve problems encountered in school mathematics. It follows closely the steps used above in mathematizing, although a key difference lies in the beginning and end of the process. Because situations in the real world are often ill-defined, the problem solver has to make some assumptions in the process of stripping away the nonessential pieces of the situation in order to arrive at a mathematical model.

In the example above, the problem solver must make some decisions about the park and the streetlight. What kind of triangle—e.g., acute, right, obtuse—best describes the park? We don't know. Will the streetlight emit light 360 degrees? We don't know. Are there structures in the park that will need to be accounted for in placing the streetlight? We don't know. Throughout the process, particularly at the beginning and the end, the problem solver must think through such aspects of the situation. A typical word problem encountered in a mathematics textbook, on the other hand, describes a situation in which some of the mathematizing has already been done. That is, some of the trimming away of the reality of the situation has been done. In these cases, what is left is the essence of the real-world situation, which already closely resembles the mathematical model that might be used to solve it.

PISA explicitly aims to measure students' ability to mathematize; thus some of the items on the PISA assessment require students to make some assumptions or ignore some information that is not relevant to the problem. NAEP items typically do not include such aspects. In NAEP, the problems that students must solve are more like those they encounter in school mathematics. The following PISA example illustrates the need for assumptions.

## M552: Rock Concert

### Question 1: ROCK CONCERT

M552Q01

For a rock concert a rectangular field of size 100 m by 50 m was reserved for the audience. The concert was completely sold out and the field was full with all the fans standing.

Which one of the following is likely to be the best estimate of the total number of people attending the concert?

A 2 000
B 5 000
C 20 000
D 50 000
E 100 000

**ROCK CONCERT SCORING 1**

***Full credit***

Code 1: C. 20 000.

***No credit***

Code 0: Other responses.

Code 9: Missing.

SOURCE: PISA 2006 released items.

From the information given in the problem, the student can calculate the size of the field to be 5,000 square meters. To convert this into the number of fans attending, the student needs to make a decision about how many fans might fit in each square meter. If the student decides that only one fan could fit in a square meter, then answer choice ‘B’ would be correct. However, the student is expected to estimate that this is not realistic for fans who are attending a rock concert, who would rather be standing close together. Taking the context into consideration, the student is expected to estimate that about 4 fans per square meter (answer choice ‘C’) is correct, after concluding that 10 (answer choice ‘D’) or 100 (answer choice ‘E’) per square meter would be too many.

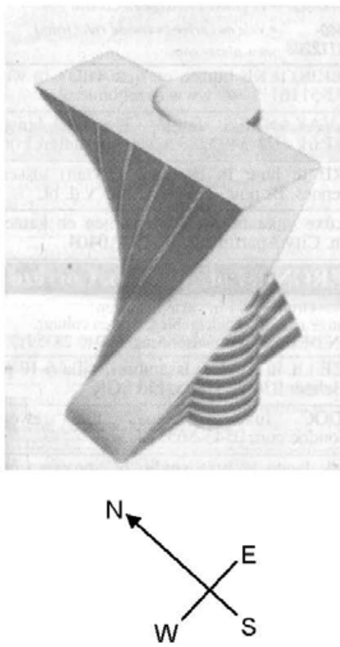
In contrast, if an item like this appeared in NAEP, it would likely include information for the student about the number of fans per square meter (4). In that case, the item would be assessing students’ ability to find the area and then understand that they need to multiply that area by 4 to judge capacity. The example PISA item is assessing an additional skill: namely, the ability to mathematize by taking this real-world situation and making appropriate assumptions about the density of the fans.

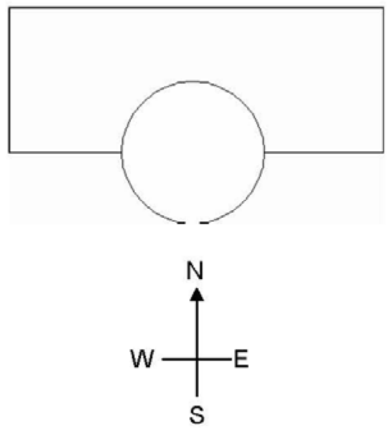
## 5.2 Role of context

As the analyses described earlier showed, the role of context differs fairly substantially between PISA and NAEP. The role of context in PISA, typically, is to set up a situation that requires some degree of mathematizing or developing a mathematical model. To solve some of the problems, students may need to make decisions or assumptions, and when this is the case, the scoring rubric will include a certain level of tolerance (i.e., a range of acceptability for such decisions). For example, in the PISA item, "Twisted Building," that follows, students must make some estimate for the height of each story of the building. The scoring rubric allows for the height to be from 2.5 meters to 4 meters, and students are expected to explain their assumptions. In this item, the context is very important to the student's reasoning.

## M535: Twisted Building

In modern architecture, buildings often have unusual shapes. The picture below shows a computer model of a 'twisted building' and a plan of the ground floor. The compass points show the orientation of the building.




The ground floor of the building contains the main entrance and has room for shops. Above the ground floor there are 20 storeys containing apartments.

The plan of each storey is similar to the plan of the ground floor, but each has a slightly different orientation from the storey below. The cylinder contains the elevator shaft and a landing on each floor.

### Question 1: TWISTED BUILDING

*M535Q01 - 0 1 2 9*

Estimate the total height of the building, in metres. Explain how you found your answer.

**TWISTED BUILDING SCORING 1**

***Full credit***

Code 2: **Accept answers from 50 to 90 metres if a correct explanation is given.**

- One floor of the building has a height of about 2.5 meters. There is some extra room between floors. Therefore an estimate is 21 x 3 = 63 metres.
- Allow 4 m for each story, so 20 of these gives 80 m, plus 10 m for the ground floor, so a total of 90 m.

Source: PISA 2006 released items.

In contrast, NAEP items that involve measurements do not require students to make assumptions. The following example is typical of a NAEP item and shows that the ranges for the dimensions of the room are given and there is only one correct answer. Given that the dimensions are prescribed, the specific context of it being a room measured by two people is irrelevant to solving the problem.

1. Carlene told Kyle that a rectangular room measured 16 feet by 12 feet, to the nearest foot. This means that the length could measure between 15.5 feet and 16.5 feet and the width could measure between 11.5 feet and 12.5 feet.

   Kyle performed the following calculations.

| Dimensions (feet) | Area (square feet) |
|---|---|
| 15 by 11 | 165 |
| 15.5 by 11.5 | 178.25 |
| 16 by 12 | 192 |
| 16.5 by 12.5 | 206.25 |
| 17 by 13 | 221 |

   Of the following intervals, which is the smallest interval that contains all possible values of the area of the room?

   A. Between 191.5 and 192.5 square feet
   B. Between 191 and 193 square feet
   C. Between 179 and 206 square feet
   D. Between 178 and 207 square feet
   E. Between 165 and 221 square feet

**Key**

D

*Source:* NAEP 2009 grade 12 released items.

To extend these analyses, the expert panel specifically reviewed those NAEP and PISA items, regardless of their context category, that were mapped to the NAEP objective solving application problems involving rational numbers and operations using exact answers or estimates as appropriate. This review revealed that there were more similarities than differences between these subsets of items. They both included some multi-step problems, calculations, and reading information from a table.

A similar analysis was done on the subset of NAEP and PISA items that were mapped to the PISA category of quantity. Again, there were similarities in numerical reasoning and using ratio and proportions. The differences were most pronounced among the NAEP items that seemed to be about quantity but did not fit the “spirit” of the PISA framework. These items included facts, straight calculations, or routine word problems. In other words, they had either no context or very minimal context.

## 5.3 Reading load

PISA items often have more text than NAEP items, which results in a heavier reading load. Some of the added text for an item (or set of items) is due to the intention to make the context more realistic; thus, information may be included that is irrelevant to solving the problem or answering the question. In some instances, additional text is needed to accommodate a translation issue across languages. Although the text of an item is generally longer in PISA than in NAEP, this does not necessarily mean that the reading difficulty of PISA items is greater than that of NAEP items.

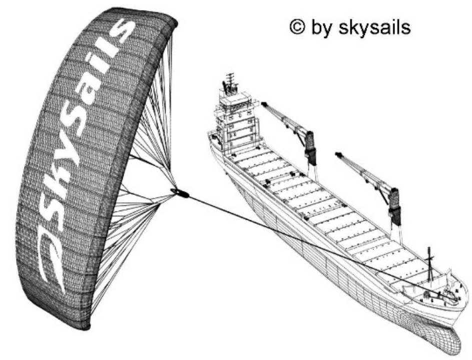
In the first example on the next page (Sailing Ships), the text describes the contextual situation for a set of PISA items. While the information enables the student to gain an understanding of the context, none of the factual information in the introductory text is essential for answering the questions that follow.

NAEP items, by contrast, generally keep the reading load to a minimum. The NAEP example following Sailing Ships exemplifies a fairly high amount of text, compared with most NAEP items. However, every piece of information in the text and diagram (with the exception of the introductory sentence) is necessary to find a solution.

# SAILING SHIPS

Ninety-five percent of world trade is moved by sea, by roughly 50 000 tankers, bulk carriers and container ships. Most of these ships use diesel fuel.

Engineers are planning to develop wind power support for ships. Their proposal is to attach kite sails to ships and use the wind's power to help reduce diesel consumption and the fuel's impact on the environment.




**Translation Note: "© by skysails": Do not adapt skysails as this is a registered label.**

## Question 1: SAILING SHIPS

*PM923Q01*

One advantage of using a kite sail is that it flies at a height of 150 m. There, the wind speed is approximately 25% higher than down on the deck of the ship.

At what approximate speed does the wind blow into a kite sail when a wind speed of 24 km/h is measured on the deck of the ship?

A 6 km/h
B 18 km/h
C 25 km/h
D 30 km/h
E 49 km/h

**Translation Note: In this unit please retain metric units throughout.**

**SAILING SHIPS SCORING 1**

QUESTION INTENT:

Description: Apply calculation of percentage within a given real world situation
Mathematical content area: Quantity
Context: Scientific
Process: Employ

***Full Credit***

Code 1: D. 30 km/h

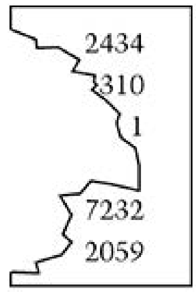
***No Credit***

Code 0: Other responses.

Code 9: Missing.

*Source:* PISA 2012 released items.

Dianne found the torn piece of paper shown below.




1. Six numbers originally appeared in a column on this paper. The fourth number from the top of the column had been completely torn away. Dianne wondered whether the sum of the six numbers was odd or even.

   Give an example of a number that could be the fourth number in the column if the sum of the six numbers is an odd number.

   Answer:_____________

   Explain why you chose that number.

*Source:* NAEP 2011 grade 8 released items.

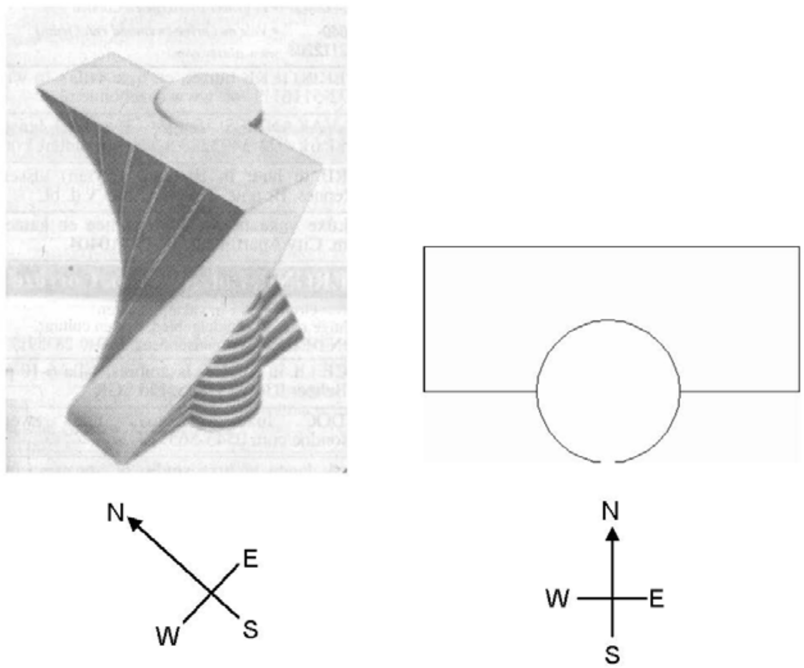
## 5.4 Use of representations

Both NAEP and PISA use a variety of representations in their mathematics items. These representations include illustrations, diagrams, tables, and graphs. However, there are differences in both the quantity and features of these representations. The stimuli for PISA items often are illustrated with photos or realistic diagrams of the problem situation, whereas the stimuli for NAEP items do not use photos and only include a diagram if it is necessary for solving the problem. When a representation is used, the stem of a NAEP item will reference it (e.g., “in the table above, …”), which is not a common practice for PISA items. Moreover, NAEP diagrams are often simpler than those found in PISA. For example, a street map on a NAEP item might be simplified to show straight lines representing streets, whereas a PISA street map would more closely resemble a “real” map, with only partial simplifications. PISA items also would more likely include graphs that are unconventional, requiring the student to make sense of a different kind of graphical representation, while NAEP graphs are restricted to a small set of easily recognizable type of graphs, such as bar graphs or line graphs.

The following two items exemplify the differences in PISA’s and NAEP’s use of illustrations and diagrams. Both items measure some form of spatial visualization. The PISA item shows a photo of a building, and the item expects students to be able to imagine the direction from which a view of the building has been drawn.

## M535: Twisted Building

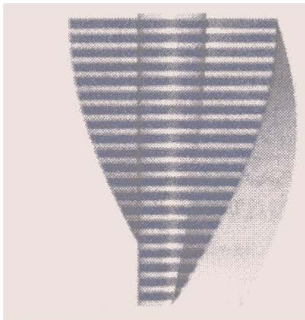
In modern architecture, buildings often have unusual shapes. The picture below shows a computer model of a 'twisted building' and a plan of the ground floor. The compass points show the orientation of the building.



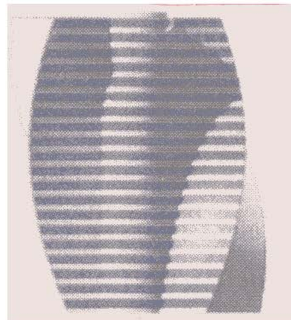

The ground floor of the building contains the main entrance and has room for shops. Above the ground floor there are 20 storeys containing apartments.

The plan of each storey is similar to the plan of the ground floor, but each has a slightly different orientation from the storey below. The cylinder contains the elevator shaft and a landing on each floor.

The following pictures are sideviews of the twisted building.



Sideview **1**



Sideview **2**

---

## Question 2: TWISTED BUILDING

*M535Q02*

From which direction has Sideview **1** been drawn?

A From the North.
B From the West.
C From the East.
D From the South.

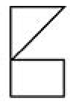
**TWISTED BUILDING SCORING 2**

***Full credit***

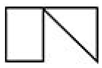
Code 1: C. From the East.

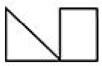
*Source:* PISA 2006 released items

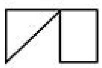
The NAEP item below depicts a drawing of a geometric figure and expects students to be able to imagine the result when the figure is rotated 90 degrees clockwise.

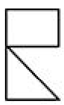


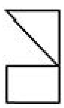
1. When the figure above is rotated 90 degrees clockwise, which of the following is the resulting figure?

A. 

B. 

C. 

D. 

E. 

**Key**

A

*Source:* NAEP 2011 grade 8 released items.

For some PISA items, an illustration is not essential to finding the answer, but may be included to enhance the translation of the context across many languages and cultures. The final example, Helen the Cyclist, shows a photo of a cyclist to provide context only, as there is no information in the photo that is needed to solve the problem. In NAEP, a similar item would be included in the assessment without the photo.

## HELEN THE CYCLIST

Helen has just got a new bike. It has a speedometer which sits on the handlebar.

The speedometer can tell Helen the distance she travels and her average speed for a trip.

**Question 1: HELEN THE CYCLIST** *PM957Q01*

On one trip, Helen rode 4 km in the first 10 minutes and then 2 km in the next 5 minutes.

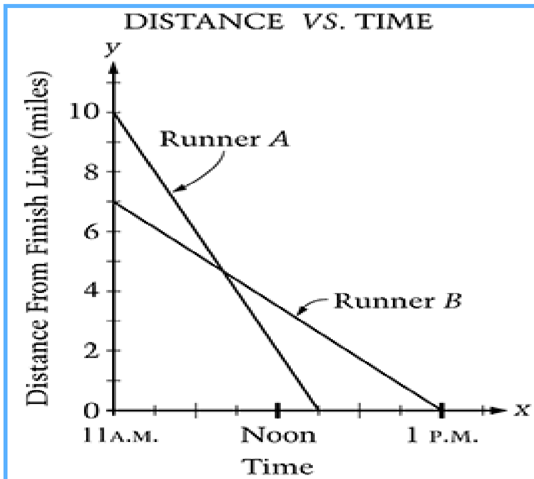
Which one of the following statements is correct?

A Helen's average speed was greater in the first 10 minutes than in the next 5 minutes.
B Helen's average speed was the same in the first 10 minutes and in the next 5 minutes.
C Helen's average speed was less in the first 10 minutes than in the next 5 minutes.
D It is not possible to tell anything about Helen's average speed from the information given.

*Source:* PISA 2012 released items.

## 5.5 Clustering and scaffolding

Both NAEP and PISA include groups of items that share stimulus material. In both assessments, these groups of items takes different forms in terms of how closely related the individual items are to each other. The main difference is that NAEP will sometimes consider each item in the group as one part of a single item with a single score. For example, in the NAEP item that follows, the item requires three responses, and the scoring rubric treats the three responses as components of a single response. Each response is scored as correct, partial, or incorrect, and then a composite score is derived from the possible combinations of scoring for the components.




The graph above shows distance versus time for a race between runners A and B. The race is already in progress, and the graph shows only the portion of the race that occurred after 11 A.M.

The table on the next page lists several features of the graph. Interpret these features in terms of what happened during this portion of the race. Include times and distances to support your interpretation. (A sample interpretation of the y-intercepts is given in the table.)

| Characteristic of Graph | Interpretation in Terms of the Race |
|---|---|
| *y*-intercepts | At 11 A.M. Runner *A* is 10 miles from the finish line and Runner *B* is 7 miles from the finish line. |
| Slopes | |
| Point of intersection | |
| *x*-intercepts | |

**Scoring Guide**

**Solution:**

There are three components in this response.

| Characteristic of Graph | Interpretation in Terms of the Race |
|---|---|
| *y*-intercepts | At 11 A.M. Runner *A* is 10 miles from the finish line and Runner *B* is 7 miles from the finish line. |
| (a) Slopes | **Runner *A*'s speed is 8 mph and Runner *B*'s speed is 3.5 mph.** |
| (b) Point of intersection | **Runner *A* and Runner *B* are both $4\frac{2}{3}$ miles from the finish line at 11:40 A.M.**<br>**NOTES:**<br>**• Accept distances from 4.5 to 5 miles, inclusive and times from 11:30 to 11:45, inclusive.**<br>**• This can also be interpreted as the time and distance when Runner *A* overtakes Runner *B*.** |
| (c) *x*-intercepts | **Runner *A* finishes the race at 12:15 P.M. and Runner *B* finishes the race at 1 P.M.** |

Note:

- Sample responses for partial credit.

| Characteristic of Graph | Interpretation Without Values | Values Only for Interpretation |
|---|---|---|
| *y*-intercepts | N/A | N/A |
| Slopes | **The speed of each runner.** | ***A* = 8 mph**<br>***B* = 3.5 mph** |
| Point of intersection | **The point where or when Runner *A* overtakes Runner *B*.**<br>**OR**<br>**The time when both runners are at the same distance from the finish line.** | **11:40AM**<br>**$4\frac{2}{3}$ miles** |
| *x*-intercepts | **The end of the race.**<br>**OR**<br>**The time that each runner finishes the race.** | ***A* = 12:15 PM**<br>***B* = 1:00 PM** |

| Score & Description |
|---|
| **Part A:**<br><br>**Correct**<br><br>Acceptable response |
| **Partial**<br><br>Acceptable interpretation without numerical values for slopes<br><br>or<br><br>Numerical values without acceptable interpretation for slopes |
| **Incorrect**<br><br>Incorrect response |
| **Part B:**<br><br>**Correct**<br><br>Acceptable response |
| **Partial**<br><br>Acceptable interpretation without numerical values for point of intersection<br><br>or<br><br>Numerical values without acceptable interpretation for point of intersection |
| **Incorrect**<br><br>Incorrect response |

*Source:* NAEP 2009 grade 12 released items.

In other cases, NAEP will use the same stimulus material, but the associated items are scored separately. PISA also has items that share the same stimulus material, but they are always treated as separate items in the scoring.

In both NAEP and PISA, groups of items that share the same stimulus material may or may not be related to each other in terms of scaffolding the student responses. Scaffolding is an intentional ordering of the items so that the student is first asked a question that may help orient the student to the problem or context; subsequent questions then build on that orientation, or even on the answer obtained in an earlier question.

Two PISA item clusters illustrate a group of independent items as well as a group of scaffolded items. In the first cluster (Climbing Mount Fuji), there is no scaffolding present: the three items are not related to each other mathematically, although they share the same stimulus material and context.

## CLIMBING MOUNT FUJI

Mount Fuji is a famous dormant volcano in Japan.

**Question 1: CLIMBING MOUNT FUJI** *PM942Q01*

Mount Fuji is only open to the public for climbing from 1 July to 27 August each year. About 200 000 people climb Mount Fuji during this time.

On average, about how many people climb Mount Fuji each day?

A 340
B 710
C 3400
D 7100
E 7400

**CLIMBING MOUNT FUJI SCORING 1**

QUESTION INTENT:

Description: Identify an average daily rate given a total number and a specific time period (dates provided)

Mathematical content area: Quantity

Context: Societal

Process: Formulate

***Full Credit***

Code 1: C. 3400

***No Credit***

Code 0: Other responses.

Code 9: Missing.

## Question 2: CLIMBING MOUNT FUJI

*PM942Q02 – 0 1 9*

The Gotemba walking trail up Mount Fuji is about 9 kilometres (km) long.

Walkers need to return from the 18 km walk by 8 pm.

Toshi estimates that he can walk up the mountain at 1.5 kilometres per hour on average, and down at twice that speed. These speeds take into account meal breaks and rest times.

Using Toshi's estimated speeds, what is the latest time he can begin his walk so that he can return by 8 pm?

**CLIMBING MOUNT FUJI SCORING 2**

QUESTION INTENT:

> Description: Calculate the start time for a trip given two different speeds, a total distance to travel and a finish time
>
> Mathematical content: Change and relationships
>
> Context: Societal
>
> Process: Formulate

***Full Credit***

Code 1: 11 (am) [with or without am, or an equivalent way of writing time, for example, 11:00]

***No Credit***

Code 0: Other responses.

Code 9: Missing.

**Question 3: CLIMBING MOUNT FUJI** *PM942Q03 – 0 1 2 9*

Toshi wore a pedometer to count his steps on his walk along the Gotemba trail.

His pedometer showed that he walked 22 500 steps on the way up.

Estimate Toshi's average step length for his walk up the 9 km Gotemba trail. Give your answer in centimetres (cm).

Answer: ................................................. cm

**CLIMBING MOUNT FUJI SCORING 3**

QUESTION INTENT:

Description: Divide a length given in km by a specific number and express the quotient in cm

Mathematical content: Quantity

Context: Societal

Process: Employ

***Full Credit***

Code 2: 40

***Partial Credit***

Code 1: Responses with the digit 4 based on incorrect conversion to centimetres.

- 0.4 *[answer given in metres]*
- 4000 *[incorrect conversion]*

***No Credit***

Code 0: Other responses.

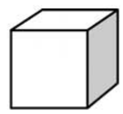
Code 9: Missing.

*Source:* PISA 2012 released items.

In the second example (Building Blocks), an introductory question asks the student to consider how many building blocks would be needed to add a layer to the one shown in the diagram. The next question asks about adding a second layer. Engaging the student in thinking about the three dimensions that are depicted in the two-dimensional diagram should enable the student to answer the next question—about imagining how many cubes are on the interior of a block—and finally to imagine building a new, hollow block with new dimensions. While none of these answers is dependent on getting a previous answer correct, the architecture of the four questions enables the student to access the mathematics at a less complex level initially and then to build toward a more complex idea. These kinds of scaffolded questions appear in both assessments, but only in a subset of the clustered items.

## M309: Building Blocks

Susan likes to build blocks from small cubes like the one shown in the following diagram:



Small cube

Susan has lots of small cubes like this one. She uses glue to join cubes together to make other blocks.

First, Susan glues eight of the cubes together to make the block shown in Diagram A:

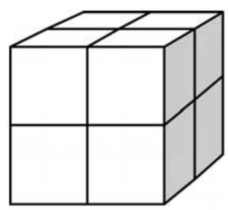


Diagram A

Then Susan makes the solid blocks shown in Diagram B and Diagram C below:

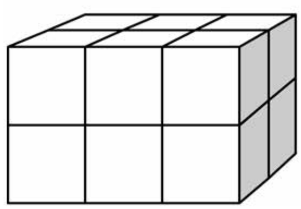


Diagram B

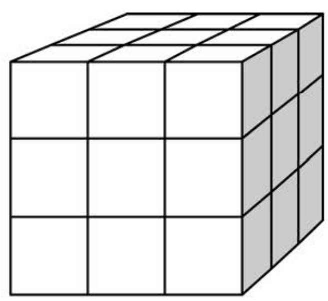


Diagram C

### Question 1: BUILDING BLOCKS

M309Q01

How many small cubes will Susan need to make the block shown in Diagram B?

Answer: ..................................................cubes.

**BUILDING BLOCKS SCORING 1**

***Full credit***

Code 1: 12 cubes.

***No credit***

Code 0: Other responses.

Code 9: Missing.

---

### Question 2: BUILDING BLOCKS

M309Q02

How many small cubes will Susan need to make the solid block shown in Diagram C?

Answer: ..................................................cubes.

**BUILDING BLOCKS SCORING 2**

***Full credit***

Code 1: 27 cubes.

***No credit***

Code 0: Other responses.

Code 9: Missing.

## Question 3: BUILDING BLOCKS

*M309Q03*

Susan realises that she used more small cubes than she really needed to make a block like the one shown in Diagram C. She realises that she could have glued small cubes together to look like Diagram C, but the block could have been hollow on the inside.

What is the minimum number of cubes she needs to make a block that looks like the one shown in Diagram C, but is hollow?

Answer: ..................................................cubes.

**BUILDING BLOCKS SCORING 3**

***Full credit***

Code 1: 26 cubes.

***No credit***

Code 0: Other responses.

Code 9: Missing.

## Question 4: BUILDING BLOCKS

*M309Q04*

Now Susan wants to make a block that looks like a solid block that is 6 small cubes long, 5 small cubes wide and 4 small cubes high. She wants to use the smallest number of cubes possible, by leaving the largest possible hollow space inside the block.

What is the minimum number of cubes Susan will need to make this block?

Answer: ..................................................cubes.

**BUILDING BLOCKS SCORING 4**

***Full credit***

Code 1: 96 cubes.

***No credit***

Code 0: Other responses.

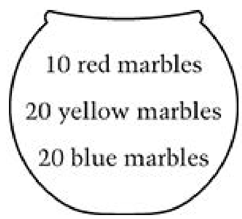
Code 9: Missing.

*Source:* PISA 2006 released items.

## 5.6 Scoring of constructed-response items

For many of the constructed-response items, the scoring guides for both NAEP and PISA are similar in that they describe responses that earn students credit. In NAEP, any item that asks for an explanation or justification will fall into at least a three-level scoring: no credit/incorrect, partial credit, and full credit/correct. In PISA, however, many of the items asking for an explanation or justification are scored dichotomously, with no partial credit awarded.

The following NAEP item asks students to decide if Juan is correct and then to explain the choice. This is scored with a three-level rubric.




1. The bowl above contains the indicated number of marbles. The marbles are well mixed in this bowl. Juan will randomly pick a marble from the bowl. Juan believes that his chance of picking a blue marble is the same as his chance of picking a yellow marble. Is Juan correct?

   Fill in the correct oval below.

   ◯ Yes ◯ No

   Explain your answer.

**Scoring Guide**

**Solution:**

Sample Correct Response:

Correct oval: Yes

Explanation:
This is true because the number of blue marbles in the bowl is equal to the number of yellow marbles.

Note(s):

- Address that the chances of picking a blue marble or picking a yellow marble are equal. It is not a complete response to restate that there are 20 blue marbles and 20 yellow marbles; it is that the numbers are equal.
- Correctly calculate the probability of selecting a blue or yellow marble (each is equal to $\frac{20}{50}$ or $\frac{2}{5}$ equivalent).

| Score & Description |
| --- |
| **Correct**<br><br>Fills in correct oval AND provides a complete explanation |
| **Partial 1** |

| |
|---|
| Fills in correct oval AND provides an incomplete or partially correct explanation |
| **Partial 2**<br><br>Fills in neither oval AND provides a complete or incomplete explanation that supports a response of "yes" |
| **Incorrect 1**<br><br>Fills in correct oval WITH incorrect explanation or no explanation |
| **Incorrect 2**<br><br>Other incorrect responses |

*Source:* NAEP 2011 grade 8 released items.

The following PISA item (Exchange Rate) also asks students to decide between two choices and then give an explanation to support their answer. The scoring rubric for this item has only two levels.

## M413: Exchange Rate

Mei-Ling from Singapore was preparing to go to South Africa for 3 months as an exchange student. She needed to change some Singapore dollars (SGD) into South African rand (ZAR).

### Question 3: EXCHANGE RATE

*M413Q03 - 01 02 11 99*

During these 3 months the exchange rate had changed from 4.2 to 4.0 ZAR per SGD.

Was it in Mei-Ling's favour that the exchange rate now was 4.0 ZAR instead of 4.2 ZAR, when she changed her South African rand back to Singapore dollars? Give an explanation to support your answer.

**EXCHANGE RATE SCORING 3**

***Full credit***

Code 11: 'Yes', with adequate explanation.

- Yes, by the lower exchange rate (for 1 SGD) Mei-Ling will get more Singapore dollars for her South African rand.
- Yes, 4.2 ZAR for one dollar would have resulted in 929 ZAR. [Note: student wrote ZAR instead of SGD, but clearly the correct calculation and comparison have been carried out and this error can be ignored]
- Yes, because she received 4.2 ZAR for 1 SGD, and now she has to pay only 4.0 ZAR to get 1 SGD.
- Yes, because it is 0.2 ZAR cheaper for every SGD.
- Yes, because when you divide by 4.2 the outcome is smaller than when you divide by 4.
- Yes, it was in her favour because if it didn't go down she would have got about $50 less.

***No credit***

Code 01: 'Yes', with no explanation or with inadequate explanation.

- Yes, a lower exchange rate is better.
- Yes it was in Mei-Ling's favour, because if the ZAR goes down, then she will have more money to exchange into SGD.
- Yes it was in Mei-Ling's favour.

Code 02: Other responses.

Code 99: Missing.

*Source:* PISA 2006 released items.

While NAEP also includes constructed-response items that are scored dichotomously, they are typically the short constructed-response items, rather than those that involve writing an explanation or justification. Additionally, some NAEP items use more than three levels of scoring, although this is not common.

# 6. Summary and Conclusion

This section summarizes the findings from across the report, drawing out some of the key findings as well as some of the implications inherent in the data.[20]

## 6.1 Framework

- *Overall, there was not a great deal of overlap between the PISA 2012 framework and the NAEP 2103 framework at grades 8 and 12.*

  Nearly all of the paired content areas were rated "quite dissimilar but with some overlap" and the paired cognitive dimensions were rated "substantially or wholly different."

- *There is a fundamental difference in the underlying approaches to assessing mathematics between the PISA 2012 framework and the NAEP framework at grades 8 and 12.*

  While this does not guarantee that the items developed from these frameworks will differ, the fact that the guidance for development differs suggests that the operationalized item pools will differ at least somewhat as well.

## 6.2 Mathematics content of item pools

- *The percentage distribution of items across content areas and categories differs between PISA and NAEP.*

  Using the NAEP framework as the basis of comparison—since this is the framework through which the full sets of items from both assessments can be compared—some differences in relative emphasis between PISA and NAEP are apparent. At grade 8, whereas NAEP has a stronger emphasis on algebra, the PISA items deemed consistent with this grade have a stronger emphasis on number properties and operations and on data analysis, statistics, and probability. At grade 12, NAEP and PISA have a similar emphasis on algebra and on data analysis, statistics, and probability, but the PISA items deemed consistent with this grade emphasize geometry more strongly than NAEP does. Moreover, there are no items in number properties and operations in PISA, in contrast to NAEP.

- *The mathematics content measured by the PISA items is most comparable to the content of the NAEP framework at grade 8.*

  The large majority of PISA items (75 percent) matched objectives in the NAEP framework at grade 8, although some items also matched objectives in the framework at grades 12 or 4. This suggests that the PISA item pool is more like the NAEP grade 8 item pool than either the NAEP grade 12 or grade 4 item pool in terms of the mathematical content being assessed.

---

[20] Note that because this section is largely focused on key cross-cutting findings, it does not repeat every finding from the report (nor does it strictly follow the order of the report itself).

- *There are some similarities in how PISA and NAEP actualize their different content areas.*
  - In both directions of the classification exercise, large percentages of all PISA items and the matched NAEP items from most content areas could be matched to the seemingly analogous category of the other's framework. For example, over 70 percent of PISA's quantity items and uncertainty and data items matched to NAEP's number properties and operations, and data analysis, statistics, and probability content areas, respectively. In the reverse direction, over 80 percent of the matched NAEP items from the geometry; algebra; and data analysis, statistics, and probability content areas at both grades matched to the corresponding category in the PISA framework. This correspondence suggests that, despite differing descriptions in the frameworks, there is overlap between PISA and NAEP in the way the individual content areas are represented in the actual assessment items.
  - The PISA items and matched NAEP items also could generally be matched to more detailed levels of the other's framework. For example, all the PISA items that matched to the NAEP framework at grade 8 or 12 (82 of 85 items) did so at the subtopic and objective level as well. Seventy-eight percent of the NAEP grade 8 items and 89 percent of the grade 12 items included in the framework comparison matched to one of the 15 topics identified across the PISA framework. This suggests that for the PISA and NAEP items that matched to the other's framework, there was a close fit to a range of the topics or subtopics in the counterpart framework.
  - Looking in particular at PISA's coverage of the subtopics in the NAEP framework, it appears that the majority of subtopics in the five content areas were covered by at least one PISA item. These are the specific areas that PISA and NAEP grade 8 and 12 items have in common in terms of content.
- *However, there also are important differences in content.*
  - PISA appears to define the change and relationships content category more broadly than NAEP does its algebra content area. PISA's items in this category matched to several different content areas in the NAEP framework, including only 52 percent to the algebra content area. In contrast, of the NAEP algebra items, 86 percent of grade 8 items and 92 percent of grade 12 items matched to PISA's change and relationships category. In the reverse direction, NAEP appears to define the measurement category in a way that spans PISA's definitions of the space and shape, and quantity categories. This suggests, as did the framework analysis, that the way individual content areas are defined differs between PISA and NAEP and differs more strongly for certain content areas (although as the prior bullets show, the "missing" content may still be covered in the other framework, just in a different content area).
  - While the PISA items covered a range of the subtopics in the NAEP framework at various grade levels, there were some subtopics they did not cover that NAEP grade 8 and 12 items did. These included what NAEP defines as mathematical reasoning using number; position, direction, and coordinate geometry; mathematical reasoning in geometry; measurement in triangles; experiments and samples; and mathematical reasoning in algebra. This analysis suggests the specific differences in the content of the PISA and NAEP item pools.

- The PISA items fit more easily to the NAEP framework than the NAEP items fit to the PISA framework. Only about half of NAEP's grade 8 and 12 items (56 and 43 percent, respectively) were considered to have a "potential fit" to PISA's framework, and of these, about one-third of grade 8 and almost one-half of grade 12 items were actually matched to the PISA framework in terms of content. (In contrast, 82 of the 85 PISA items matched to the NAEP framework at either grade 8 or grade 12.) This suggests that while we might find the content of PISA items on a NAEP assessment, we would be less likely to find the content of NAEP items on a PISA assessment.

## 6.3 Cognitive dimension of item pools

- The fit of PISA and NAEP items to each other's framework in terms of the cognitive dimension was uniformly strong, compared to the fit to the content dimensions. With one exception each, all of the PISA items and all of the NAEP grade 8 and 12 items in the classification exercise could be categorized in the counterpart framework: in NAEP's mathematical complexity levels and PISA's mathematical processes. This is not surprising, as the cognitive dimension for both frameworks is generally defined more as a continuum into which it is likely that almost any item could fit.
- In terms of mathematical complexity, PISA items tended to have fewer items in the low category than NAEP did overall (44 percent compared to 56 percent at both grades) and more items in the moderate category (53 percent compared to 39 percent at both grades). Neither assessment had very large percentages of items in the high category (just 2 percent of PISA items and 5 percent of NAEP grade 8 and grade 12 items). However, in the subset of the NAEP item pool that was deemed to match the PISA framework in terms of content, the distributions were much more favorable to higher complexity levels—that is, the small number of NAEP items that were most PISA-like were generally of higher complexity than the NAEP and PISA items as a whole.
- In terms of mathematical processes, PISA tended to have fewer items than NAEP did in the more straightforward category of "employ" (44 percent compared to more than 50 percent in NAEP at both grades) and more in the "formulate" category (32 percent compared to no more than 27 percent at both grades). PISA had more than twice the percentage of items as NAEP did in the "interpret" category (25 percent compared to 9 and 12 percent at both grade 8 and grade 12). However, these comparisons are based on only those NAEP items included in the classification exercise, and so it is unknown what the distribution of the full NAEP item pools would be.

## 6.4 Context of item pools

- The PISA framework calls for all PISA items to be placed in a context. In contrast, although NAEP does specify that some items should have a focus on real-world problem solving, it is not required of all items.

- Not surprisingly, just one PISA item (1 percent of all items) was judged to have no context. In comparison, 6 and 11 percent of the NAEP items at grades 8 and 12 included in the classification exercise were not in a specific context and were, rather, straightforward "naked" mathematics items. These percentages would likely be higher for NAEP if the entire item pool was included, considering that lack of context was a primary reason for the exclusion of items from the classification exercise.
- At the other end of the spectrum, PISA items were about twice as likely as the classified NAEP items to be context dependent (31 percent compared to 14 and 17 percent at grades 8 and 12, respectively).

## 6.5 Item features

- PISA and NAEP items differ in the extent to which students are required to *mathematize* problem situations. PISA explicitly aims to measure students' ability to mathematize; thus, some of the items on the PISA assessment require students to make some assumptions or ignore some information that is not relevant to the problem. NAEP items typically do not include such aspects.
- The *role of context* in PISA is often integral to the requirement for students to mathematize (and thus integral to the functioning of the items), whereas in NAEP, context generally does not play this role and is not as integral to the functioning of the items.
- PISA items often have more text than NAEP items, which indicates a heavier *reading load*. Some of the added text in a PISA item (or set of items) is due to the intention to make the context more realistic by including information that is irrelevant to solving the problem or answering the question. At times, additional text is needed to accommodate a translation issue across languages. In NAEP, reading loads are generally kept to a minimum and generally all of the information provided in the item is relevant to solving the problem.
- Both NAEP and PISA use a variety of *representations* in their mathematics items, but there are differences in both the quantity and features of these representations. In NAEP, representations are more likely to be simplified (e.g., a street map converted into simple lines) and to be confined to a specified list of graphs (e.g., bar graphs and box plots). PISA, on the other hand, is more likely to include photos, realistic diagrams, and unconventional graphs.
- Both NAEP and PISA use a set of stimulus material that is linked with multiple items—i.e., there are *clustered items* in both assessments. However, NAEP will more frequently score the associated items collectively, whereas PISA always treats the associated items separately for scoring. Both NAEP and PISA use *scaffolding*—that is, intentionally building on a previous item—in some but not all of their clustered items.
- NAEP items are more likely than PISA items to use multi-level *scoring* (e.g., full credit, partial credit, incorrect) for constructed-response items, whereas PISA items are more likely to be scored dichotomously (e.g., correct/incorrect).

In summary, the comparisons between PISA and NAEP, while showing some areas of correspondence, also reveal some key differences in both the intentions and implementation of the mathematics assessments. These differences include those in the mathematics content covered, grade-level correspondence, cognitive skills required, extent of contextualization of the items, and other item features.

All of these factors together may result in differences in reported student performance, and it is useful to consider these differences when interpreting the results from the assessments. They also are useful in understanding the unique contribution of each assessment program to the understanding of U.S. students' mathematics performance.

# References

Common Core State Standards Initiative. (2012). *Standards for Mathematical Practice.* Retrieved from http://www.corestandards.org/Math

Gattis, K., Rosenberg, S., Neidorf, T., Zhang, J., Guile, S., and McNeely, M. (2013). *A Comparison of the 2011 Grade 8 NAEP and TIMSS Mathematics and Science Frameworks*). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/nationsreportcard/about/international.aspx

National Assessment Governing Board (NAGB). (2007). *Assessment and Item Specifications for the NAEP 2009 Mathematics Assessment.* Washington, DC: Author.

National Assessment Governing Board (NAGB). (2012). *Mathematics Framework for the 2013 National Assessment of Educational Progress.* Washington, DC: Author.

National Center for Education Statistics (NCES). (2007). *Comparing PIRLS and PISA with NAEP in Reading, Mathematics, and Science*. Retrieved from http://nces.ed.gov/surveys/PISA/pdf/comppaper12082004.pdf

National Center for Education Statistics (NCES). (2008). *Comparing TIMSS with NAEP and PISA in Mathematics and Science.* Retrieved from http://nces.ed.gov/timss/pdf/Comparing_TIMSS_NAEP_%20PISA.pdf

National Center for Education Statistics (NCES). (2013). NAEP Questions Tool [data tool]. Available from http://nces.ed.gov/nationsreportcard/itmrlsx/search.aspx?subject=mathematics

Neidorf, T. S., Binkley, M., Gattis, K., & Nohara, D. (2006). *Comparing Mathematics Content in the National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), and Program for International Student Assessment (PISA) 2003 Assessments* (NCES 2006-029). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Niss, M. (2003). Mathematical competencies and the learning of mathematics: The Danish KOM project. In A. Gagatsis and S. Papastavridis (Eds.), *3rd Mediterranean Conference on Mathematics Education* (pp. 115–124). Athens, Greece: The Hellenic Mathematical Society and Cyprus Mathematical Society. Retrieved from *http://w3.msi.vxu.se/users/hso/aaa_niss.pdf*

Nohara, D. H. (2001). *Comparison of National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)* (NCES 2001-07). Washington, DC: National Center for Education Statistics.

Organization for Economic Cooperation and Development (OECD). (2003). *PISA 2003 Assessment Framework for Mathematics, Reading, Science, and Problem Solving Knowledge and Skills*. Paris: Author.

Organization for Economic Cooperation and Development (OECD). (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy.* Paris: OECD Publishing. Retrieved from *http://dx.doi.org/10.1787/9789264190511-en*

Polya, G. (1945). *How to Solve It: A New Aspect of Mathematical Method.* Princeton, NJ: Princeton University Press.

Provasnik, S., Lin, C., Darling, D., and Dodson, J., (2013). *A Comparison of the 2011 Trends in International Mathematics and Science Study (TIMSS) Assessment Items and the 2011 National Assessment of Educational Progress (NAEP) Frameworks*. U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/nationsreportcard/about/international.aspx

# Appendix A. NCES Comparison Studies of National and International Mathematics Assessments

| Author | Assessments in the comparison | Features of the comparison |
|---|---|---|
| Past studies | | |
| Nohara 2001 (working paper) | PISA 2000<br>NAEP 2000<br>TIMSS 1999 | • Classification of PISA and TIMSS items to NAEP framework in terms of content<br>• Classification of all items according to external criteria related to thinking skills and schematic related to difficulty factors |
| Neidorf et al. 2006 (technical report) | PISA 2003<br>NAEP 2003<br>TIMSS 2003 | • Classification of PISA and TIMSS items to NAEP framework in terms of content and complexity<br>• Classification of NAEP and TIMSS problem-solving items to PISA framework in terms of competencies and context |
| NCES 2007 (briefing paper) | PISA 2006<br>NAEP 2005/7 | • No new analyses<br>• Use of 2003 classifications of PISA items to NAEP framework in content and to PISA framework in competencies because 2006 items were a subset of 2003 |
| NCES 2008 (briefing paper) | PISA 2006<br>NAEP 2007<br>TIMSS 2007 | • Classification of TIMSS items to NAEP framework in terms of content and complexity<br>• Use of 2003 classifications of PISA items to NAEP framework because 2006 items were a subset of 2003 |
| Provasnik et al. 2013 (working paper) | NAEP 2011<br>TIMSS 2011 | • Classification of TIMSS items to NAEP framework in terms of content |
| Gattis et al. 2013 | NAEP 2011<br>TIMSS 2011 | • Comparison of assessment frameworks |
| Current study | | |
| Gattis et al. 2016 (working paper) | PISA 2012<br>NAEP 2013 | • Classification of PISA items to NAEP framework in terms of content and complexity<br>• Classification of NAEP items to PISA framework in terms of content and mathematical processes<br>• Classification of PISA and NAEP items to external schema on the role of context<br>• Quantitative approach to framework analysis<br>• Comparison of item attributes |

# Appendix B. Sample Data Collection Instrument

**PISA NAEP Framework Comparison** Rater ____________________

**How similar or different are the knowledge and skills specified in the NAEP framework to the knowledge and skills specified in the PISA framework?**

**4** - Exactly or almost the same **3** - Quite similar, but with some differences **2** - Quite dissimilar, but with some overlap **1** - Substantially or wholly different

| NAEP Framework | PISA Framework | Similarity Rating | | | |
|---|---|---|---|---|---|
| **Content Dimension** | | **First Individual** | | **Final Individual** | |
| **Group 1: Number Subtopics** | **Quantity** | Grd 8 | Grd 12 | Grd 8 | Grd 12 |
| Number Sense | Quantity | | | | |
| Estimation | Quantity | | | | |
| Number operations | Quantity | | | | |
| Ratios and proportional reasoning | Quantity | | | | |
| Properties of number and operations | Quantity | | | | |
| Mathematical reasoning using number | Quantity | | | | |
| **Overall Content Area: Number Properties and Operations** | **Overall Content Category: Quantity** | | | | |
| **Group 2: Measurement Subtopics** | **Space and Shape** | Grd 8 | Grd 12 | Grd 8 | Grd 12 |
| Measuring physical attributes | Space and Shape | | | | |
| Systems of measurement | Space and Shape | | | | |
| Measurement in triangles | Space and Shape | | | | |
| **Overall Content Area: Measurement** | **Overall Content Category: Space and Shape** | | | | |
| **Group 3: Geometry Subtopics** | | Grd 8 | Grd 12 | Grd 8 | Grd 12 |
| Dimension and shape | Space and shape | | | | |
| Transformation of shapes and preservation of properties | Space and shape | | | | |
| Relationships between geometric figures | Space and shape | | | | |
| Position, direction and coordinate geometry | Space and shape | | | | |
| Mathematics reasoning in geometry | Space and shape | | | | |
| **Overall Content Area: Geometry** | **Overall Content Category: Space and Shape** | | | | |
| **Overall Content Areas: Measurement and Geometry** | **Overall Content Category: Space and Shape** | | | | |
| **Group 4: Data Analysis, Statistics and Probability Subtopics** | **Uncertainty and Data** | Grd 8 | Grd 12 | Grd 8 | Grd 12 |
| Data representation | Uncertainty and Data | | | | |
| Characteristics of data sets | Uncertainty and Data | | | | |
| Experiments and samples | Uncertainty and Data | | | | |
| Probability | Uncertainty and Data | | | | |
| Mathematical reasoning with data | Uncertainty and Data | | | | |
| **Overall Content Area: Data Analysis, Statistics and Probability** | **Overall Content Category: Uncertainty and Data** | | | | |
| **Group 5: Algebra subtopics** | **Change and Relationships** | Grd 8 | Grd 12 | Grd 8 | Grd 12 |
| Patterns, relations and functions | Change and Relationships | | | | |
| Algebraic representations | Change and Relationships | | | | |
| Variables, expressions and operations | Change and Relationships | | | | |
| Equations and inequalities | Change and Relationships | | | | |
| Mathematical reasoning in algebra | Change and Relationships | | | | |
| **Overall Content Area: Algebra** | **Overall Content Category: Change and Relationships** | | | | |
| **Group 6: Overall Content Dimension** | | Grd 8 | Grd 12 | Grd 8 | Grd 12 |
| **All NAEP Content Areas** | **All PISA Content Categories** | | | | |
| **Group 7: Cognitive Dimension** | | | | | |
| **Levels of Complexity** | **Mathematical Processes** | | | | |
| Low | Formulating situations mathematically | | | | |
| Moderate | Employing mathematical concepts, facts, procedures and reasoning | | | | |
| High | Interpreting, applying and evaluating mathematical outcomes | | | | |
| | **Mathematical Capabilities** | | | | |
| | Communication | | | | |
| | Reasoning | | | | |
| | Representation | | | | |
| | Reasoning and argument | | | | |
| | Devising strategies for solving problems | | | | |
| | Using symbolic, formal and technical language and operations | | | | |
| | Using mathematical tools | Grd 8 | Grd 12 | Grd 8 | Grd 12 |
| **Overall Cognitive Dimension** | **Overall Cognitive Dimension** | | | | |
| **Group 8: Context** | | | | | |
| | Personal | | | | |
| | Occupational | | | | |
| | Societal | | | | |
| | Scientific | Grd 8 | Grd 12 | Grd 8 | Grd 12 |
| | **Overall Context Dimension** | | | | |
| **Group 9: Overall Frameworks** | | Grd 8 | Grd 12 | Grd 8 | Grd 12 |
| **Overall NAEP Framework** | **Overall PISA Framework** | | | | |

# Appendix C. Analysis of the PISA 2012 Computer-Based Items

PISA 2012 included a computer-based assessment of mathematics, which was optional for participating countries, given their varied technological capacities. There are two aspects to the rationale behind PISA's inclusion of computer-based assessment items. First, as defined in the 21st century, mathematical literacy requires at least some knowledge of how to use a computer. Second, the computer provides a range of opportunities for designers to write test items that are more interactive, authentic, and engaging.

The PISA computer-based items were classified in the same way as the paper-and-pencil items (i.e., to the NAEP framework's content and mathematical processes dimensions as well as to context categories developed for this study). Unlike the classification of the paper-and pencil items, however, the computer-based items were classified only by two mathematics experts, not all four members of the expert panel. A match was determined by agreement between the two experts.

Appendix C includes results tables for the computer-based items in terms of content, mathematical complexity, and context classifications. In addition, it includes a table for item format distribution.

**Table C-1. Number (N) and percentage (%) distribution of PISA computer-based items across NAEP content areas, by PISA content category**

| PISA content category | Grade | Any NAEP content area | | Number properties and operations | | Geometry | | Measurement | | Data analysis, statistics, and probability | | Algebra | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | N | % | N | % | N | % | N | % | N | % | N | % |
| Change and relation-ships[1] | No match | 0 | 0 | † | † | † | † | † | † | † | † | † | † |
| | Match | 11 | 100 | 1 | 9 | 0 | 0 | 0 | 0 | 2 | 18 | 8 | 73 |
| | Grade 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Grade 8 | 7 | 64 | 1 | 9 | 0 | 0 | 0 | 0 | 1 | 9 | 5 | 45 |
| | Grade 12 | 4 | 36 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 9 | 3 | 27 |
| Space and shape[2] | No match | 6 | 50 | † | † | † | † | † | † | † | † | † | † |
| | Match | 6 | 50 | 0 | 0 | 4 | 33 | 2 | 17 | 0 | 0 | 0 | 0 |
| | Grade 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Grade 8 | 6 | 50 | 0 | 0 | 4 | 33 | 2 | 17 | 0 | 0 | 0 | 0 |
| | Grade 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Uncertainty and data[3] | No match | 1 | 11 | † | † | † | † | † | † | † | † | † | † |
| | Match | 8 | 89 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 89 | 0 | 0 |
| | Grade 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Grade 8 | 8 | 89 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 89 | 0 | 0 |
| | Grade 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Quantity[3] | No match | 0 | 0 | † | † | † | † | † | † | † | † | † | † |
| | Match | 9 | 100 | 4 | 44 | 0 | 0 | 5 | 56 | 0 | 0 | 0 | 0 |
| | Grade 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Grade 8 | 9 | 100 | 4 | 44 | 0 | 0 | 5 | 56 | 0 | 0 | 0 | 0 |
| | Grade 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

[1] Total number of items mapped in this category was 11.
[2] Total number of items mapped in this category was 12.
[3] Total number of items mapped in this category was 9.
† Not applicable.
*Note:* N = Number; % = Percentage. Detail may not sum to totals because of rounding.

**Table C-2a. Number and percentage distribution of the PISA computer-based items matched to the NAEP framework at grade 8 and the NAEP grade 8 items, by NAEP content area**

| | PISA | | NAEP | |
|---|---|---|---|---|
| **NAEP content area** | **Number** | **Percent** | **Number** | **Percent** |
| Number properties and operations | 5 | 17 | 29 | 19 |
| Geometry | 4 | 13 | 26 | 17 |
| Measurement | 7 | 23 | 29 | 19 |
| Data analysis, statistics, and probability | 9 | 30 | 23 | 15 |
| Algebra | 5 | 17 | 46 | 30 |

*Note:* Detail may not sum to totals because of rounding.

**Table C-2b. Number and percentage distribution of the PISA computer-based items matched to the NAEP framework at grade 12 and the NAEP grade 12 items, by NAEP content area**

| | PISA | | NAEP | |
|---|---|---|---|---|
| **NAEP content area** | **Number** | **Percent** | **Number** | **Percent** |
| Number properties and operations | 0 | 0 | 22 | 12 |
| Geometry | 0 | 0 | 37 | 19 |
| Measurement | 0 | 0 | 22 | 12 |
| Data analysis, statistics, and probability | 1 | 25 | 48 | 25 |
| Algebra | 3 | 75 | 62 | 32 |

*Note:* Detail may not sum to totals because of rounding.

**Table C-3a. Number and percentage distribution of the PISA computer-based items and the 86 NAEP grade 8 items classified to the PISA framework, by PISA content category**

| PISA content category | PISA | | NAEP | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Change and relationships | 11 | 27 | 12 | 14 |
| Space and shape | 12 | 29 | 6 | 7 |
| Uncertainty and data | 9 | 22 | 6 | 7 |
| Quantity | 9 | 22 | 5 | 6 |
| Not matched | † | † | 57 | 66 |

† Not applicable. All PISA items can be matched to the content categories in the PISA framework.
*Note:* Detail may not sum to totals because of rounding.

**Table C-3b. Number and percentage distribution of the PISA computer-based items and the 83 NAEP grade 12 items classified to the PISA framework, by PISA content category**

| PISA content category | PISA | | NAEP | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Change and relationships | 11 | 27 | 13 | 16 |
| Space and shape | 12 | 29 | 8 | 10 |
| Uncertainty and data | 9 | 22 | 13 | 16 |
| Quantity | 9 | 22 | 4 | 5 |
| Not matched | † | † | 45 | 54 |

† Not applicable. All PISA items can be matched to the content categories in the PISA framework.
*Note:* Detail may not sum to totals because of rounding.

**Table C-4. Number and percentage distribution of PISA computer-based items, by NAEP complexity level and PISA content category**

| PISA content category | Total number of items | NAEP complexity level | | | | | | Not matched | |
|---|---|---|---|---|---|---|---|---|---|
| | | Low | | Moderate | | High | | | |
| | | Number | Percent | Number | Percent | Number | Percent | Number | Percent |
| All categories | 41 | 25 | 61 | 14 | 34 | 2 | 5 | 0 | 0 |
| Change and relationships | 11 | 8 | 73 | 3 | 27 | 0 | 0 | 0 | 0 |
| Space and shape | 12 | 4 | 33 | 6 | 50 | 2 | 17 | 0 | 0 |
| Uncertainty and data | 9 | 6 | 67 | 3 | 33 | 0 | 0 | 0 | 0 |
| Quantity | 9 | 7 | 78 | 2 | 22 | 0 | 0 | 0 | 0 |

*Note:* Detail may not sum to totals because of rounding.

**Table C-5a. Percentage distribution of PISA computer-based items and NAEP grade 8 items, by NAEP mathematical complexity level and PISA content category**

| PISA content category | Low | | Moderate | | High | | Not matched | |
|---|---|---|---|---|---|---|---|---|
| | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) |
| All NAEP items | 61 | 56 | 34 | 39 | 5 | 5 | 0 | † |
| NAEP classified only | 61 | 24 | 34 | 66 | 5 | 10 | 0 | † |
| Change and relationships | 73 | 42 | 27 | 42 | 0 | 17 | 0 | † |
| Space and shape | 33 | 33 | 50 | 67 | 17 | 0 | 0 | † |
| Uncertainty and data | 67 | 0 | 33 | 83 | 0 | 17 | 0 | † |
| Quantity | 78 | 0 | 22 | 100 | 0 | 0 | 0 | † |

† Not applicable.
*Note:* Detail may not sum to totals because of rounding.

**Table C-5b. Percentage distribution of PISA computer-based items and NAEP grade 12 items, by NAEP mathematical complexity level and PISA content category**

| PISA content category | Low | | Moderate | | High | | Not matched | |
|---|---|---|---|---|---|---|---|---|
| | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) | PISA (%) | NAEP (%) |
| All NAEP items | 61 | 56 | 34 | 39 | 5 | 5 | 0 | † |
| NAEP classified only | 61 | 34 | 34 | 58 | 5 | 8 | 0 | † |
| Change and relationships | 73 | 54 | 27 | 46 | 0 | 0 | 0 | † |
| Space and shape | 33 | 0 | 50 | 75 | 17 | 25 | 0 | † |
| Uncertainty and data | 67 | 46 | 33 | 54 | 0 | 0 | 0 | † |
| Quantity | 78 | 0 | 22 | 75 | 0 | 25 | 0 | † |

† Not applicable.
*Note:* Detail may not sum to totals because of rounding.

**Table C-6a. Number and percentage distribution of PISA computer-based items and the 86 NAEP grade 8 items classified to the PISA framework, by PISA mathematical process**

| PISA mathematical process | PISA | | NAEP | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Employ | 22 | 54 | 57 | 66 |
| Formulate | 9 | 22 | 20 | 23 |
| Interpret | 10 | 24 | 8 | 9 |
| Not matched | † | † | 1 | 1 |

† Not applicable. All PISA items could be categorized within the PISA mathematical processes.
*Note:* Detail may not sum to totals because of rounding.

**Table C-6b. Number and percentage distribution of PISA computer-based items and the 83 NAEP grade 12 items classified to the PISA framework, by PISA mathematical process**

| PISA mathematical process | PISA | | NAEP | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Employ | 22 | 54 | 51 | 61 |
| Formulate | 9 | 22 | 22 | 27 |
| Interpret | 10 | 24 | 10 | 12 |
| Not matched | † | † | 0 | 0 |

† Not applicable. All PISA items could be categorized within the PISA mathematical processes.
*Note:* Detail may not sum to totals because of rounding.

**Table C-7. Number and percentage distribution of the PISA 2012 computer-based items and the classified NAEP 2013 items, by context category**

| Context | PISA | | NAEP grade 8 | | NAEP grade 12 | |
|---|---|---|---|---|---|---|
| | Number | Percent | Number | Percent | Number | Percent |
| Not classified | † | † | 67 | 44 | 108 | 57 |
| Classified | 41 | 100 | 86 | 100 | 83 | 100 |
| No context[1] | 4 | 10 | 5 | 6 | 9 | 11 |
| Contextualized[1] | 31 | 76 | 69 | 80 | 60 | 72 |
| Context dependent[1] | 6 | 15 | 12 | 14 | 14 | 17 |

†Not applicable. All PISA items were classified to a context category.
[1]Percentage based on classified items only.
*Note:* Detail may not sum to totals because of rounding.

**Table C-8. Number and percentage distribution of PISA 2012 computer based items, by PISA content category and context category**

| Context | Change and relationships | | Space and shape | | Uncertainty and data | | Quantity | |
|---|---|---|---|---|---|---|---|---|
| | Number | Percent | Number | Percent | Number | Percent | Number | Percent |
| All items | 11 | 100 | 12 | 100 | 9 | 100 | 9 | 100 |
| No context | 0 | 0 | 4 | 33 | 0 | 0 | 0 | 0 |
| Contextualized | 10 | 91 | 6 | 50 | 8 | 89 | 7 | 78 |
| Context dependent | 1 | 9 | 2 | 17 | 1 | 11 | 2 | 22 |

*Note:* Detail may not sum to totals because of rounding.

**Table C-9. Item format descriptions and percentage distribution of PISA 2012 computer-based items and NAEP 2013 items, by format**

| Item format (PISA name/NAEP name) | Description | Percentage of PISA items | Percentage of NAEP grade 8 items | Percentage of NAEP grade 12 items |
|---|---|---|---|---|
| Selected response/ Multiple choice | Requires student to select from a number of response options | 37 | 75 | 69 |
| Closed constructed response/ Short constructed response | Requires student to generate a response (e.g., provide a number or a category) that can sometimes be scored automatically | 54 | 22 | 25 |
| Open constructed response/ Extended constructed response | Requires student to generate a longer response, often requiring an explanation or justification, that must be scored manually | 10 | 5 | 13 |

*Note:* Detail may not sum to totals because of rounding.

# Appendix D. Analysis of Content Matching of PISA 2012 and NAEP 2013 Items to NAEP Subtopics

The NAEP mathematics framework divides mathematics into five broad content areas with several major subtopics in each area. For example, two of the six subtopics in the number properties and operations content area are “number sense” and “estimation.” For each subtopic, the framework defines a set of grade-specific objectives that specify the particular knowledge and skills that students are expected to know and be able to do. Assessment items are then written to measure part or all of each objective. Both the number of subtopics and objectives vary greatly across content areas.

Coverage of NAEP mathematics subtopics in the PISA 2012 and NAEP 2013 assessments was defined as follows: A *subtopic* was determined to be covered if at least one objective in the subtopic was covered in the assessment. An *objective* was determined to be covered if at least three panel members decided that an item matched with an objective, with a fit rating of at least 3 (i.e., “fits quite well”).

**Exhibit D-1. Coverage of NAEP mathematics subtopics in the PISA 2012 and NAEP 2013 assessments, by grade, content area, and subtopic**

| Content area | Subtopic | PISA coverage of NAEP subtopics at grade | | | NAEP coverage of NAEP subtopics at grade | | |
|---|---|---|---|---|---|---|---|
| | | 4 | 8 | 12 | 4 | 8 | 12 |
| Number properties and operations | 1) Number sense | x | x | | x | x | x |
| | 2) Estimation | | | | x | x | x |
| | 3) Number operations | x | x | | x | x | x |
| | 4) Ratios and proportional reasoning | | x | | x | x | x |
| | 5) Properties of number and operations | | x | | x | x | x |
| | 6) Mathematical reasoning using number | | | | x | | x |
| Geometry | 1) Dimension and shape | | x | x | x | x | x |
| | 2) Transformation of shapes and preservation of properties | | x | | x | x | x |
| | 3) Relationships between geometric figures | | x | x | x | x | x |
| | 4) Position, direction, and coordinate geometry | | | | x | x | x |
| | 5) Mathematical reasoning in geometry | | | | x | x | x |
| Measurement | 1) Measuring physical attributes | | x | | x | x | x |
| | 2) Systems of measurement | | x | | x | x | x |
| | 3) Measurement in triangles | † | | | † | x | x |
| Data analysis, statistics, and probability | 1) Data representation | x | x | | x | x | x |
| | 2) Features of data sets | | x | | x | x | x |
| | 3) Experiments and samples | † | | | † | x | x |
| | 4) Probability | | x | x | x | x | x |
| | 5) Mathematical reasoning with data | † | † | | † | † | x |
| Algebra | 1) Patterns, relations, and functions | | x | | x | x | x |
| | 2) Algebraic representations | | | x | x | x | x |
| | 3) Variables, expressions, and operations | | x | | x | x | x |
| | 4) Equations and inequalities | | x | x | x | x | x |
| | 5) Mathematical reasoning in algebra | | | | x | | x |

† Not applicable. The subtopic is not included in the NAEP framework at the specified grade level and thus a comparison with PISA at this grade level and subtopic is not applicable.
*Note:* A subtopic was treated as covered if there was at least one item that was intended to measure at least one objective in the subtopic.

# Appendix E. Analysis of Classification of NAEP 2013 Items to PISA 2012 Content Topics

In addition to the four mathematical content categories, the PISA mathematics framework lists 15 content topics that reflect the mathematics that students around the world have likely had the opportunity to learn by the time they are 15 years old. The list is intended to be illustrative of the content topics included in PISA 2012, rather than exhaustive.

Each panelist classified the NAEP items into one of the 15 PISA content topics. These content topics may span more than one content category. The final determination of a match was made based on an agreement of at least three panelists with a fit rating of at least 3 (i.e., "fits quite well").

**Table E-1a. Number (N) and percentage (%) of NAEP grade 8 items classified to PISA content topics, by NAEP content area and classification status**

| Classification status | All items classified | | Classified items by content area | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Number properties and operations | | Geometry | | Measurement | | Data analysis, statistics, and probability | | Algebra | |
| | N | % | N | % | N | % | N | % | N | % | N | % |
| Total | 86 | 100 | 15 | 100 | 8 | 100 | 21 | 100 | 22 | 100 | 20 | 100 |
| No agreement | 16 | 19 | 5 | 33 | 1 | 13 | 4 | 19 | 3 | 14 | 3 | 15 |
| Agreement with specific PISA topics | 67 | 78 | 7 | 47 | 7 | 88 | 17 | 81 | 19 | 86 | 17 | 85 |
| Agreement with no PISA topics | 3 | 3 | 3 | 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

*Note:* N = Number; % = Percentage. Detail may not sum to totals because of rounding.

**Table E-1b. Number (N) and percentage (%) of NAEP grade 12 items classified to PISA content topics, by NAEP content area and classification status**

| Classification status | Classified items by content area | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | All items classified | | Number properties and operations | | Geometry | | Measurement | | Data analysis, statistics, and probability | | Algebra | |
| | N | % | N | % | N | % | N | % | N | % | N | % |
| Total | 83 | 100 | 8 | 100 | 7 | 100 | 12 | 100 | 41 | 100 | 15 | 100 |
| No agreement | 9 | 11 | 3 | 38 | 0 | 0 | 0 | 0 | 2 | 5 | 4 | 27 |
| Agreement with specific PISA topics | 74 | 89 | 5 | 63 | 7 | 100 | 12 | 100 | 39 | 95 | 11 | 73 |
| Agreement with no PISA topics | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

*Note:* N = Number; % = Percentage. Detail may not sum to totals because of rounding.

**Table E-2a. Number (N) and percentage (%) of NAEP grade 8 items matched to a specific PISA content topic, by PISA topic**

| | Total | | Number properties and operations | | Geometry | | Measurement | | Data analysis, statistics, and probability | | Algebra | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **PISA topic** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** | **N** | **%** |
| Total | 67 | 100 | 7 | 100 | 7 | 100 | 17 | 100 | 19 | 100 | 17 | 100 |
| Functions | 7 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 41 |
| Algebraic expressions | 5 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 29 |
| Equations and inequalities | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Coordinate systems | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 6 |
| Relationships within and among geometrical objects in two and three dimensions | 5 | 7 | 0 | 0 | 5 | 71 | 0 | 0 | 0 | 0 | 0 | 0 |
| Measurement | 8 | 12 | 0 | 0 | 0 | 0 | 8 | 47 | 0 | 0 | 0 | 0 |
| Numbers and units | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 6 | 0 | 0 | 0 | 0 |
| Arithmetic operations | 1 | 1 | 1 | 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Percents, ratios, and proportions | 17 | 25 | 5 | 71 | 2 | 29 | 8 | 47 | 1 | 5 | 1 | 6 |
| Counting principles | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Estimation | 1 | 1 | 1 | 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Data collection, representation, and interpretation | 7 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 21 | 3 | 18 |
| Data variability and its description | 5 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 26 | 0 | 0 |
| Samples and sampling | 3 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 16 | 0 | 0 |
| Chance and probability | 6 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 32 | 0 | 0 |

*Note:* N = Number; % = Percentage. Detail may not sum to totals because of rounding.

**Table E-2b. Number (N) and percentage (%) of NAEP grade 12 items matched to a specific PISA content topic, by PISA topic**

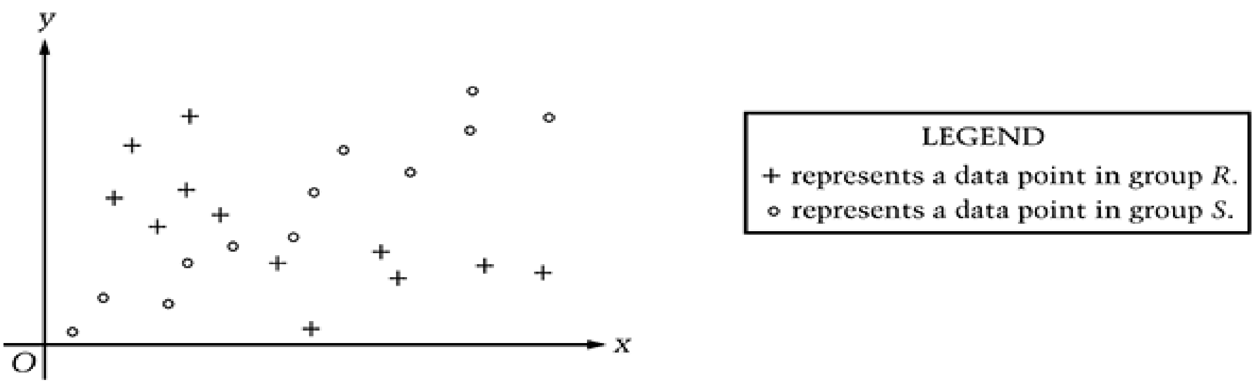
| PISA topic | Total | | Number properties and operations | | Geometry | | Measurement | | Data analysis, statistics, and probability | | Algebra | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | % | N | % | N | % | N | % | N | % | N | % |
| Total | 74 | 100 | 5 | 100 | 7 | 100 | 12 | 100 | 39 | 100 | 11 | 100 |
| Functions | 6 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 55 |
| Algebraic expressions | 5 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 5 | 3 | 27 |
| Equations and inequalities | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 9 |
| Coordinate systems | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Relationships within and among geometrical objects in two and three dimensions | 6 | 8 | 0 | 0 | 4 | 57 | 2 | 17 | 0 | 0 | 0 | 0 |
| Measurement | 8 | 11 | 0 | 0 | 3 | 43 | 5 | 42 | 0 | 0 | 0 | 0 |
| Numbers and units | 2 | 3 | 1 | 20 | 0 | 0 | 0 | 0 | 1 | 3 | 0 | 0 |
| Arithmetic operations | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Percents, ratios, and proportions | 13 | 18 | 4 | 80 | 0 | 0 | 4 | 33 | 5 | 13 | 0 | 0 |
| Counting principles | 2 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 5 | 0 | 0 |
| Estimation | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 8 | 0 | 0 | 0 | 0 |
| Data collection, representation, and interpretation | 11 | 15 | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 26 | 1 | 9 |
| Data variability and its description | 8 | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 21 | 0 | 0 |
| Samples and sampling | 2 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 5 | 0 | 0 |
| Chance and probability | 9 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 9 | 23 | 0 | 0 |

*Note:* N = Number; % = Percentage. Detail may not sum to totals because of rounding.

# Appendix F. Context Categories Illustrated Through NAEP and PISA Items

**Items Classified as No Context**

*This example is a NAEP grade 12 data analysis and probability item; note that the graph and variables are not assigned any contextual references.*



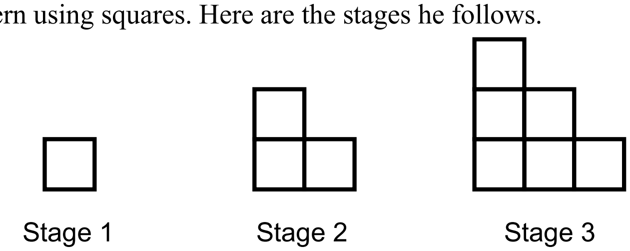

The scatterplot above shows data for groups R and S . Which of the following statements is true about the correlation between the x and y values of group R and the correlation between the x and y values of group S?

A. The x and y values appear to be negatively correlated in both R and S .
B. The x and y values appear to be positively correlated in both R and S .
C. The x and y values appear to be negatively correlated in R , but positively correlated in S .
D. The x and y values appear to be positively correlated in R , but negatively correlated in S .
E. The x and y values appear to be more highly correlated in R than in S .

*Source:* NAEP 2009 Grade 12 Released Items.

*This example is a PISA quantity item; note that there are geometric figures involved, which are considered purely mathematical.*

Robert builds a step pattern using squares. Here are the stages he follows.




As you can see, he uses one square for Stage 1, three squares for Stage 2 and six for Stage 3.

How many squares should he use for the fourth stage?

*Source:* PISA 2003 Released Items.

## Items Classified as Contextualized

*This example is a NAEP 2011 grade 8 number item, which demonstrates the minimal end of the contextualized spectrum. Note that there is a context, but that it is not required for the interpretation of either part of the item.*

| Item | Cost |
|---|---|
| Yogurt | $0.95 each |
| Pretzels | $2.50 per bag |
| Cheese cubes | $2.19 per bag |
| Bagel | $0.89 each |
| Fruit drink | $1.85 each |
| Peanuts | $2.55 per bag |

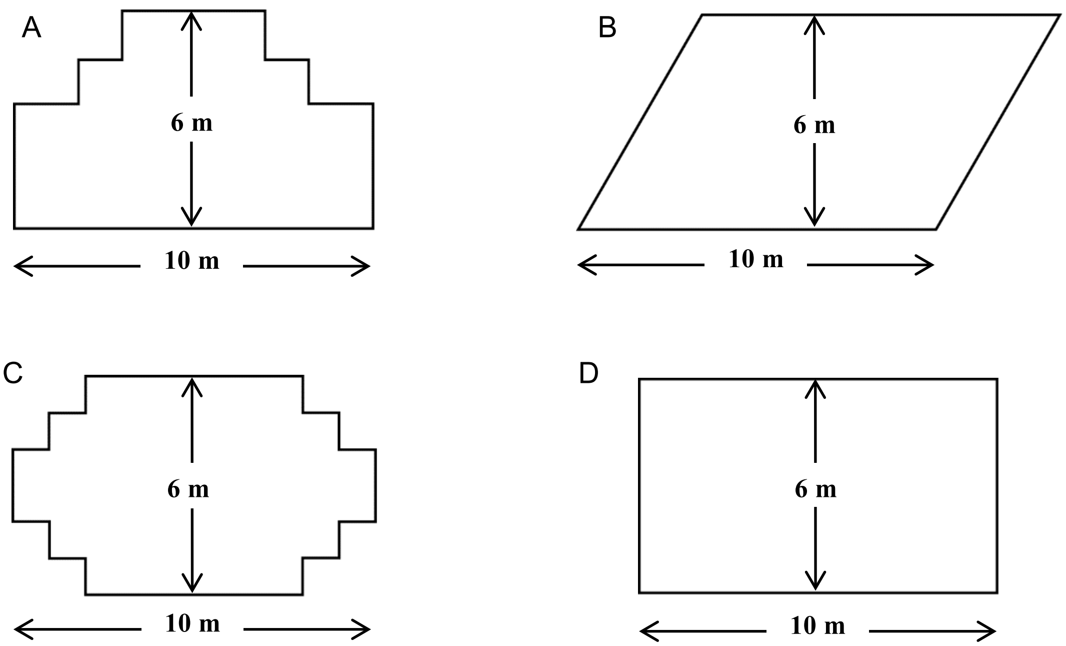
Robert has $30 and wants to buy as many bags of peanuts as possible. He does not have to pay any sales tax on the food that he buys.

(a) Based on the prices given in the chart above, how many bags of peanuts can Robert buy?
Answer: ____________________
(b) Robert buys all the bags of peanuts that he can. What is the most expensive single item on the chart that he can buy with the money he has left?
Answer: ____________________

*Source:* NAEP 2011 Grade 8 Released Items.

*This example is a PISA space and shape item. Note that while there is a context, the translation from the situation described in the problem into the geometric figure has already happened and thus the student could answer the questions with minimal attention to the context.*

A carpenter has 32 meters of timber and wants to make a border around a garden bed. He is considering the following designs for the garden bed.



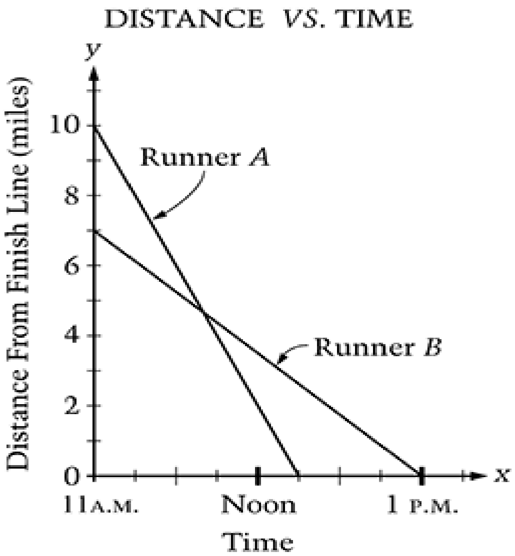

Circle either "Yes" or "No" for each design to indicate whether the garden bed can be made with 32 meters of timber.

| **Garden bed design** | **Using this design, can the garden bed be made with 32 meters of timber?** |
|---|---|
| Design A | Yes / No |
| Design B | Yes / No |
| Design C | Yes / No |
| Design D | Yes / No |

*Source:* PISA 2012 Released Items.

## Items Classified as Context-Dependent

*This example is a NAEP 2009 grade 12 algebra item. Note that the context must be used to interpret aspects of the graph, and the student's response requires that the context be taken into account.*




The graph above shows distance versus time for a race between runners A and B. The race is already in progress, and the graph shows only the portion of the race that occurred after 11 A.M.

The table on the next page lists several features of the graph. Interpret these features in terms of what happened during this portion of the race. Include times and distances to support your interpretation. (A sample interpretation of the y-intercepts is given in the table.)

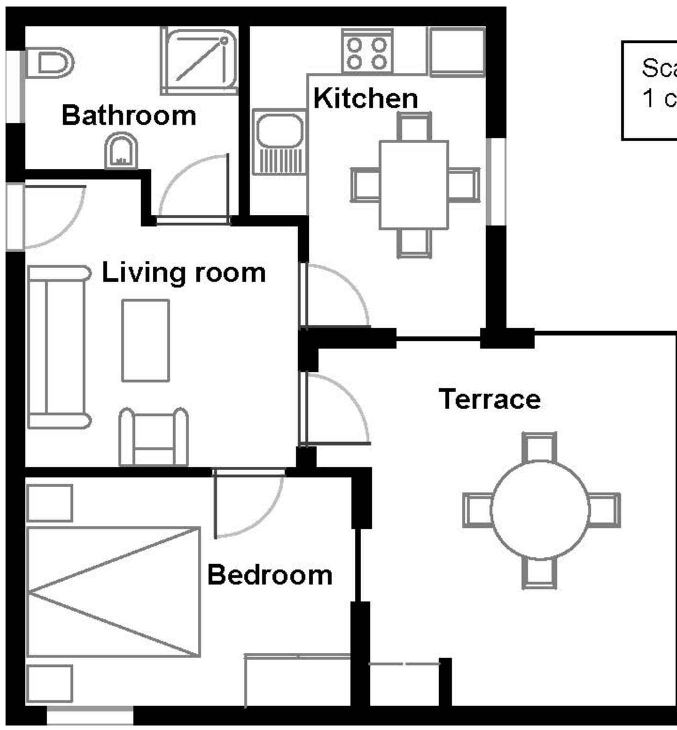
| Characteristic of Graph | Interpretation in Terms of the Race |
|---|---|
| *y*-intercepts | At 11 A.M. Runner *A* is 10 miles from the finish line and Runner *B* is 7 miles from the finish line. |
| Slopes | |
| Point of intersection | |
| *x*-intercepts | |

*Source:* NAEP 2009 Grade 12 Released Items.

*This example is a PISA 2003 space and shape item. Note that the item requires a stripping away of all the features of the floor plan except for the outside walls, recognizing that none of the other details of the plan (such as doorways, interior walls, or windows) are relevant to a solution.*

# APARTMENT PURCHASE

This is the plan of the apartment that George's parents want to purchase from a real estate agency.




Scale:
1 cm represents 1 m

## Question 1: APARTMENT PURCHASE

*PM00FQ01 – 0 1 9*

To estimate the total floor area of the apartment (including the terrace and the walls), you can measure the size of each room, calculate the area of each one and add all the areas together.

However, there is a more efficient method to estimate the total floor area where you only need to measure 4 lengths. Mark on the plan above the **four** lengths that are needed to estimate the total floor area of the apartment.

*Source:* PISA 2003 Released Items.

## ABOUT AMERICAN INSTITUTES FOR RESEARCH

Established in 1946, with headquarters in Washington, D.C., American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally. As one of the largest behavioral and social science research organizations in the world, AIR is committed to empowering communities and institutions with innovative solutions to the most critical challenges in education, health, workforce, and international development.


AIR®

AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

**www.air.org**


*Making Research Relevant*

## LOCATIONS

### Domestic

Washington, D.C.

Atlanta, GA

Austin, TX

Baltimore, MD

Cayce, SC

Chapel Hill, NC

Chicago, IL

Columbus, OH

Frederick, MD

Honolulu, HI

Indianapolis, IN

Metairie, LA

Naperville, IL

New York, NY

Rockville, MD

Sacramento, CA

San Mateo, CA

Waltham, MA

### International

Egypt

Honduras

Ivory Coast

Kyrgyzstan

Liberia

Tajikistan

Zambia


02-2016_04/16