5. NAEP/TIMSS Comparisons by Main Content Areas

The overall comparisons section highlighted that NAEP and TIMSS have similar emphases across the broad mathematics content areas. This section provides more detailed comparisons of the content coverage of the items for each of the NAEP and TIMSS mathematics content area subscales. There are five content area sections presenting comparisons for items in each of the NAEP *content strands* with those in the corresponding TIMSS *content domains*. These five content-area sections include:[1]

- *number*;
- *measurement*;
- *geometry*;
- *data*; and
- *algebra*.

Each content area section includes

- a comparison of the relevant parts of the content frameworks;[2]
- an analysis of the level of match between the items from one assessment and the topics and subtopics included at particular grades in the other assessment framework; and
- a comparison of how items are distributed across topics within these content areas as defined by each framework.

For these analyses, the NAEP and TIMSS items are divided according to subscale and then comparisons are made within the content areas that are the same or similar across the two assessments. Content and grade classification are examined simultaneously in the analyses for this section. For each content area, the report reports the percentage of items that were classified to the other framework at the corresponding grade level or at another grade level. For items classified at the corresponding grade level, there are three levels of content match, as follows:

- specific match (to a specific subtopic in the same content area);[3]
- general match (at the broader topic level in the same content area but not the subtopic level); and

[1] Although NAEP and TIMSS use somewhat different labels to refer to each of these content areas, the one-word TIMSS categories are used in the discussions in this section for the sake of convenience.
[2] Framework comparison tables in this section list the topics included in the content area. Additional information about the specific subtopics included for each of the main topics is given in appendix A.
[3] Specific match also applies to items classified to a topic in the NAEP framework for which no subtopics are included.

- match to another content area (at either the topic or subtopic level).

For items classified at another grade level in the other assessment framework, there are two types, as follows:

- lower grade (grade 8 items classified to grade 4 topics or subtopics) and

- higher grade (grade 4 items classified to grade 8 topics or subtopics or grade 8 TIMSS items classified to grade 12 NAEP topics or subtopics)[4]

This section also reports the percentages of items not classified to topics in the other assessment framework (i.e., those that could not be classified to a topic or subtopic at a specific grade).[5] The text in this section may refer to items classified to specific subtopics, but this level of detail is not shown in the tables in the report. Subtopics are shown in appendix A and example items illustrating various features that are referenced in this section are shown in appendix E.

5.1. Number

The *number* content area receives greater emphasis at fourth grade than at eighth grade for both the NAEP and TIMSS assessments. As discussed in section 4, in NAEP, 42 percent of fourth-grade items and 26 percent of eighth-grade items are from the *content strand* of *number sense, properties, and operations* (table 4). In TIMSS, 38 percent of fourth-grade items and 31 percent of eighth-grade items are from the *content domain* of *number* (table 5). The results in the *number* section are based on 76 fourth-grade and 51 eighth-grade items in NAEP, and 55 fourth-grade and 56 eighth-grade items in TIMSS.

Framework comparison in number

The framework structures used in NAEP and TIMSS to organize content in this area are quite different. Exhibit 3 shows a comparison of the *number* topics included in the NAEP and TIMSS mathematics frameworks. The NAEP framework is organized around skills, such as *represent numbers and operations in a variety of equivalent forms using models, diagrams and symbols*, and *compute with numbers*. The TIMSS framework, on the other hand, is organized by types of numbers, for example, *whole numbers*, and *fractions and decimals*. When considering all subtopics (appendix A), there are fewer differences between the content specified in NAEP and TIMSS. NAEP subtopics often specify types of numbers and TIMSS subtopics specify types of skills. The correspondence is not complete, however, leaving some ambiguity regarding whether a topic or subtopic not mentioned specifically in one framework might still be implied. For example, one TIMSS subtopic is "solve problems with fractions." Similar content might be implied in the NAEP subtopic of "solve application problems involving numbers and operations, using exact answers or estimates, as appropriate," although fractions are not mentioned specifically.

[4] Because the NAEP 2003 framework is used to guide a twelfth-grade assessment and the TIMSS 2003 framework is not, the classification of grade 8 items to the twelfth-grade level is only applicable for the classification of TIMSS items to the NAEP framework.

[5] The method for determining grade-level match in this section differs somewhat from what was used for the overall comparisons in section 4.2. Overall comparisons of grade level include items classified at any level of content match (subtopic, topic, or broad content area). In this section, grade level was not assigned unless items could be classified to at least the topic level.

One difference resulting from the different organizational approaches is that the NAEP framework contains more detail about the aspects of computation to be emphasized. Whereas both frameworks address computation, in TIMSS it is at the subtopic level, under two separate topics, *whole numbers* and *fractions and decimals*. In NAEP, it is at the topic level (*compute with numbers*) and is further specified by four subtopics. One thing noted by the panel related to computations was that the NAEP framework does not explicitly include non-contextualized computation that is not placed in a problem-solving or application setting. Although there is a main topic in NAEP entitled *compute with numbers*, the subtopic of "apply basic properties of operations" was interpreted by the panel as requiring an application context. In contrast, in the TIMSS framework, there are explicit subtopics related to the mechanics and properties of computation within the main topics by number type (e.g., "compute with whole numbers;" and "add and subtract fractions or decimals"). This difference is primarily an issue with the specificity and interpretation of the framework, however, since both assessments include computation items.

One similarity between the frameworks is that both focus on representation and computation involving whole numbers, fractions, and decimals at the fourth-grade level, with ratios, proportions, and percents and working with integers not emphasized until the eighth-grade level.[6]

Exhibit 3. Number topics included in the NAEP and TIMSS mathematics frameworks: 2003

NAEP ***Number sense, properties, and operations topics***	**TIMSS** ***Number topics***
Relate counting, grouping and place value	Whole numbers
Represent numbers and operations in a variety of equivalent forms using models, diagrams and symbols	Fractions and decimals
Compute with numbers (that is, add, subtract, multiply, divide)	Integers (grade 8 only)
Use computations and estimation in applications	Ratio, proportion, and percent
Apply ratios and proportional thinking in a variety of situations	
Use elementary number theory	

NOTE: Unless otherwise noted, all topics are intended for all grades (grades 4, 8, and 12 in NAEP and grades 4 and 8 in TIMSS). The number of subtopics or objectives at each grade and level of detail varies across topics and assessments.
SOURCE: U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.*, 2003.

Content and grade match in number

With some exceptions, NAEP and TIMSS appear to share similar definitions of the content area of *number* at both the fourth and eighth grades, since, for each assessment most items were classified at either the topic or subtopic level on the framework of the other assessment at the same grade level (table 10). In NAEP, 79 percent of fourth-grade and 73 percent of eighth-grade items were classified at the same grade level with either a specific or general match to TIMSS *number* topics. There was an even higher correspondence between the TIMSS items and *number* topics in

[6] Although the word *integer* does not appear in the NAEP framework topics, it does appear in the grade-level item illustrations given in the assessment specifications document (NAGB 1992).

the NAEP framework, with 98 percent of fourth-grade and 96 percent of eighth-grade items having a specific or general match to *number* topics in the NAEP framework.

The 22 percent of TIMSS fourth-grade items with a general match were all classified to the NAEP topic of *compute with numbers*. Panelists noted that these items were of two types, which they described as, "basic computation" and "select an operation or procedure to solve a problem." Examples of each of these types of TIMSS items are shown in Examples 10 and 11, respectively. As defined by the panel, "basic computation" items were items that required students to perform a computation in a non-contextualized situation and involved only the mechanics and properties of computation. Panelists also identified several items of these types on the TIMSS eighth-grade assessment as well. As noted previously, the panel interpreted the subtopic within the NAEP *compute with numbers* topic as requiring an application context, so any purely computational items in TIMSS were classified at only the general topic level in the NAEP framework.

An examination of the classification of NAEP items to TIMSS subtopics (data not shown) revealed that NAEP also includes computation items, although there were not as many of these types of items as in TIMSS. The TIMSS framework includes three fourth-grade subtopics within the topics of *whole numbers* and *fractions and decimals* that address only computation. Although relatively more TIMSS items were placed in these subtopics, some NAEP items were also classified to the computation subtopics in TIMSS. Almost all of these NAEP items came from a single NAEP subtopic, *apply basic properties of operations*. The appearance of the computation items in NAEP indicates that the strict interpretation of this subtopic by the panel to mean that an application setting is required was not applied in the same fashion by the NAEP assessment developers. An item from the NAEP assessment that was described by the panel as "basic computation" is shown in Example 12 in appendix E.

At both grade levels, all or almost all TIMSS items were classified at the same grade level on the NAEP framework (table 10). On the other hand, 16 percent of fourth-grade NAEP items were classified as eighth-grade items on the TIMSS framework, almost all as *fractions and decimals* spread across various subtopics. An example of a NAEP fourth-grade *number* item placed at the eighth-grade level on the TIMSS mathematics framework is illustrated by Example 13. Twenty-seven percent of eighth-grade NAEP items were placed at the fourth-grade level on the TIMSS framework, all but one of which were cross-grade items administered at both fourth and eighth grades. Most of these items were placed in the TIMSS topics of *whole numbers* and *fractions and decimals* and came from several different NAEP topics and subtopics.

Although all but 4 percent of eighth-grade TIMSS items were placed at the eighth-grade level on the NAEP framework, based on additional comments made by the panel roughly one third of TIMSS items were noted as being slightly below the specifications for the eighth grade. Therefore, while these items were judged to correspond most closely with the general descriptions of the eighth-grade level in the NAEP framework, they were found to have some characteristics consistent with a somewhat lower level than eighth grade (but not consistent with the specifications at the fourth-grade level).

A few eighth-grade NAEP items were identified that included the use of scientific notation, which was not addressed in the TIMSS framework subtopics, nor reflected in the TIMSS items (data not shown). Also, there were a number of NAEP items that include mathematical operations involving money. Some of these items had a general match to TIMSS, since they were consistent with the descriptions of the broader topics. Some items in the fourth grade that require knowledge of the value of specific U.S. coins, however, were classified as not having a match in the TIMSS framework topics, as items of this type would not be included in an international assessment like TIMSS. In TIMSS and PISA, any items with money contexts use a common fictitious currency such as *zeds* (see example 7 in appendix E).

Table 10. Percentage of NAEP and TIMSS fourth- and eighth-grade number items classified to the other mathematics assessment framework, by level of content/grade match: 2003

Level of content/grade match	Grade 4		Grade 8	
	NAEP items to TIMSS framework	TIMSS items to NAEP framework	NAEP items to TIMSS framework	TIMSS items to NAEP framework
Total number of number items	76	55	51	56
	Percentage distribution			
Classified as same grade	79	100	73	96
Specific match[1] in number	71	76	47	80
General match[2] in number	8	22	25	16
Match to another content area[3]	0	2	0	0
Classified as another grade[4]	16	0	27	4
Lower grade[5]	†	†	27	4
Higher grade[6]	16	0	†	0
No classification to topics[7]	5	0	0	0

† Not applicable. Grade 4 is the lowest grade in both frameworks; grade 8 is the highest grade in the TIMSS 2003 framework.
[1] Includes items that were classified at the subtopic level at the same grade (and items classified to NAEP framework topics for which no subtopics are included).
[2] Includes items that were classified to a topic but not to a subtopic at the same grade.
[3] Includes items that were classified to a topic or a subtopic in a different content area at the same grade.
[4] Includes items that were classified to a topic or subtopic in any content area at another grade.
[5] Includes grade 8 items classified to grade 4 topics/subtopics.
[6] Includes grade 4 items classified to grade 8 topics/subtopics or grade 8 TIMSS items classified to grade 12 NAEP topics/subtopics.
[7] Includes items that the panel did not classify to a topic at a specific grade level.
NOTE: Data reflect the percentage of items classified by the expert panel at each level. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.*, 2003.

Item distribution across number topics

Figures 8 and 9 show the percentage of NAEP and TIMSS number items across the topics included in the *number* frameworks for each assessment. The distribution of items across NAEP topics indicates a greater emphasis in NAEP on the application of computation (as opposed to the mechanics and properties of computation) than in TIMSS. At the fourth-grade level, whereas TIMSS includes a higher percentage of items classified to the NAEP framework as *compute with numbers* (27 percent compared to 20 percent of NAEP items), NAEP includes a higher percentage of items classified as *use computations and estimation in applications* (33 percent compared to 24 percent). This pattern is found at the eighth-grade level as well—8 percent of NAEP items compared to 25

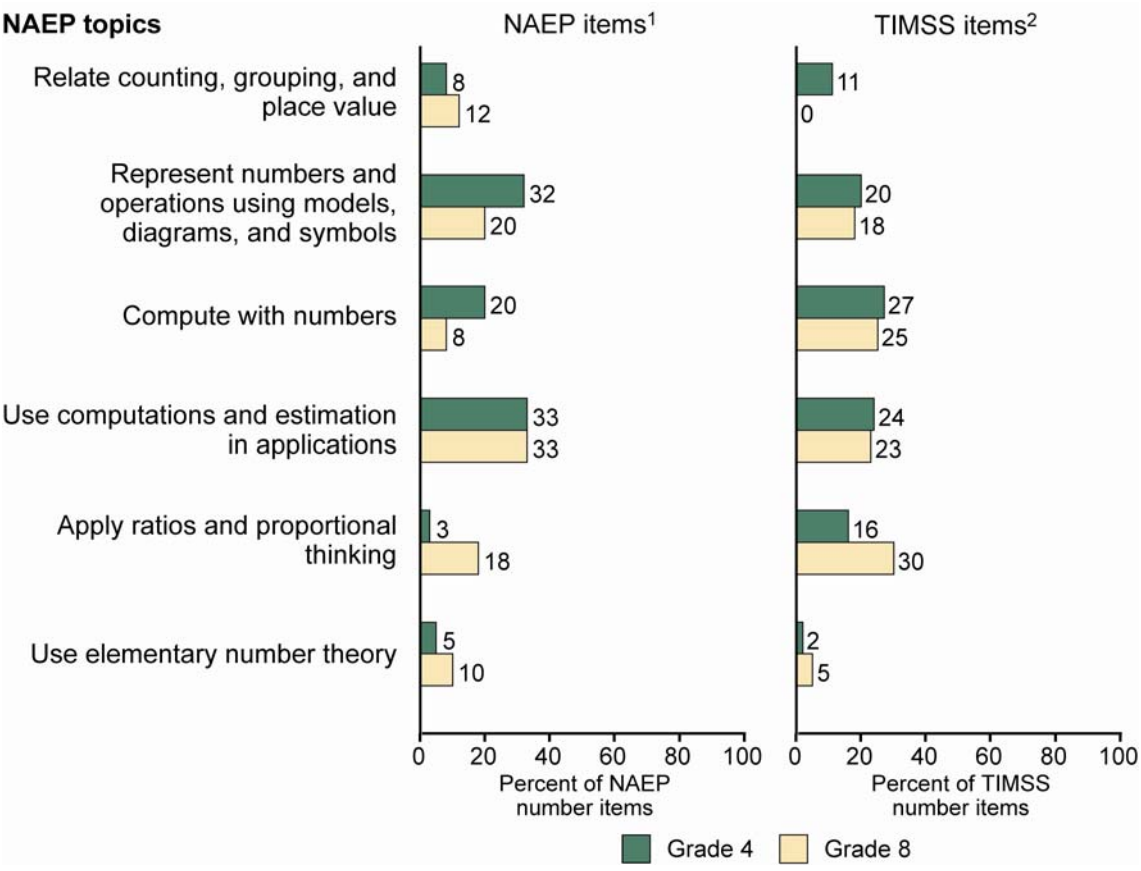
percent of TIMSS items classified to the *compute with numbers* topic and 33 percent of NAEP items compared to 23 percent of TIMSS items classified to the NAEP topic of *use computations and estimation in applications*. However, this finding is contradicted by the fact that almost half of all items on both eighth-grade assessments were classified to TIMSS subtopics that dealt with problem solving across the main topic areas, such as solving problems involving decimals, fractions, integers, etc. (data not shown). This points to some lack of specificity in both frameworks in distinguishing between pure computation and application or problem-solving contexts.

Another difference is that TIMSS includes a higher percentage of items involving ratios and proportions. This is supported by classifications on both frameworks—the NAEP topic of *apply ratios and proportional thinking* and the TIMSS topic of *ratio, proportion, and percent*. Considering both classification systems, TIMSS has 13 and 16 percent, respectively, of items in this topic area at the fourth grade compared to 3 and 5 percent in NAEP at the fourth grade. At eighth grade, TIMSS has about 30 percent of items focused on *ratio, proportion, and percent* compared to about 18 percent in NAEP. The NAEP eighth-grade assessment has a higher proportion of items classified as *whole numbers* on the TIMSS framework (37 percent compared to 14 percent of TIMSS).

Other notable differences based on the NAEP framework include a higher percentage of fourth-grade NAEP items classified as *represent numbers and operations using models, diagrams, and symbols* (32 percent compared to 20 percent in TIMSS). Also, at the eighth grade, 12 percent of NAEP items were classified to the topic of *relate counting, grouping, and place value* (half of which addressed scientific notation); in TIMSS there were items classified to this topic at fourth grade but none at eighth grade, and no items involving scientific notation.

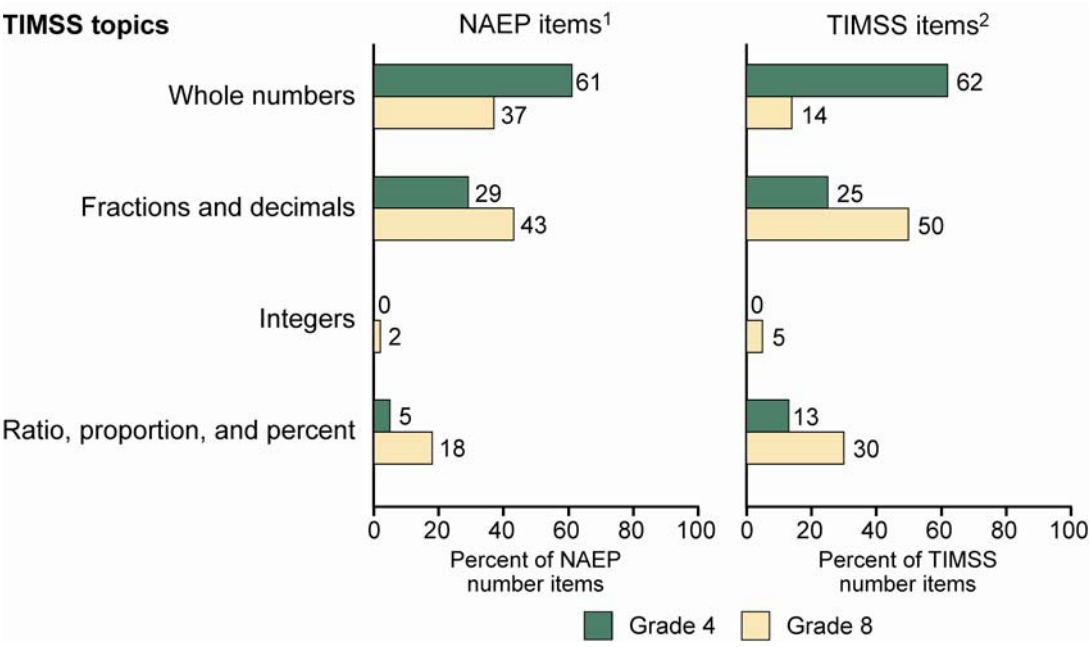
In both NAEP and TIMSS, the fourth-grade assessments emphasize items involving *whole numbers* (more than 60 percent of items), while the eighth-grade assessments are focused on items involving *fractions and decimals* (43 and 50 percent of items, respectively). A small percentage of items (5 percent or less) involving *integers* were in included in both assessments at the eighth grade but not the fourth grade.

Figure 8. Percentage of NAEP and TIMSS number items classified to number sense, properties, and operations topics in the NAEP mathematics framework, by survey and grade: 2003



[1] NAEP items classified by NAEP developers.
[2] TIMSS items classified by expert panel.
NOTE: Topics may be abbreviated for graphical clarity. Percentages reflect the proportion of *number* items classified at either the topic level or the subtopic level at any grade level. Items that were classified to multiple topics were counted in all relevant topics. Bars not shown indicate that no items from that particular grade and assessment were classified to the topic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; and U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002.

Figure 9. Percentage of NAEP and TIMSS number items classified to number topics in the TIMSS mathematics framework, by survey and grade: 2003



[1] NAEP items classified by expert panel.
[2] TIMSS items classified by TIMSS developers.
NOTE: Percentages reflect the proportion of *number* items classified at either the topic level or the subtopic level at any grade level. Items that were classified to multiple topics were counted in all relevant topics. Bars not shown indicate that no items from that particular grade and assessment were classified to the topic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.*, 2003.

5.2. Measurement

Measurement items contribute similar proportions of the NAEP and TIMSS assessments at both grade levels. As discussed in section 4, in NAEP, 18 percent of fourth-grade items and 15 percent of eighth-grade items are from the *content strand* of *measurement* (table 4). In TIMSS, 21 percent of fourth-grade items and 16 percent of eighth-grade items are from the *content domain* of *measurement* (table 5). The results in the *measurement* section are based on 32 fourth-grade and 30 eighth-grade items in NAEP, and 31 fourth-grade and 28 eighth-grade items in TIMSS.

Framework comparison in measurement

Similar to the *number* frameworks, the NAEP and TIMSS *measurement* frameworks appear quite different at the topic level. However, when all subtopics included in NAEP and TIMSS are considered (appendix A), there is considerable overlap. Exhibit 4 shows a comparison of the *measurement* topics included in the NAEP and TIMSS mathematics frameworks. The TIMSS framework includes two broad topics with a set of eight subtopics across these two topics at both the fourth- and eighth-grade levels. The NAEP framework includes 10 topics, only three of which include subtopics. In addition, only five of the NAEP topics are intended to be included in the fourth-grade assessment. Comparing the topics and subtopics at a general level, both frameworks appear to address the major content areas typically associated with measurement, including measuring and computing attributes of figures (e.g., length, area, volume, perimeter); selecting appropriate tools, units, and methods; and converting units of measure.

Content and grade match in measurement

The level of content and grade match for NAEP and TIMSS *measurement* items is shown in table 11. Fourth-grade NAEP *measurement* items had a closer match to TIMSS *measurement* topics and subtopics than fourth-grade TIMSS items did to NAEP, with 85 percent of NAEP items having either a specific or general match compared to 74 percent of TIMSS items. One reason for this difference is that the fourth-grade TIMSS assessment contained a number of items (16 percent) classified by the panel as being consistent with the eighth-grade NAEP framework. Most of these were nevertheless classified as *measurement* items.

In contrast, at the eighth-grade level, there was a closer match between the TIMSS items and the NAEP framework than between the NAEP items and the TIMSS framework. More than two-thirds of TIMSS items had either a specific or general match to the NAEP framework, whereas less than half of NAEP items did so to the TIMSS framework. A substantial number of items on both the TIMSS and NAEP eighth-grade assessments (25 and 37 percent) were classified as having a better match to the descriptions at the fourth-grade level of the other assessment framework. Example 14 in appendix E illustrates a TIMSS eighth-grade *measurement* item placed at the fourth-grade level on the NAEP mathematics framework.

NAEP contained a larger percentage of items that, although classified at the corresponding grade level, were classified to another content area (13 percent of fourth-grade and 17 percent of eighth-grade items in NAEP compared to about 5 percent of TIMSS items). Most of these NAEP items involved measurements of angles or properties of geometric shapes and were classified to *geometry* topics in TIMSS.

In *measurement* as well as in *number*, there was a case of similar items appearing on both assessments but apparently serving different purposes. Not classified to an appropriate NAEP subtopic were several TIMSS fourth-grade *measurement* items that required students to perform calculations with time and temperature. The NAEP fourth-grade assessment does include items of this type, but they were classified across various NAEP topics and subtopics, meaning that they served a purpose other than assessing students' ability to perform calculations with these types of measures.

Exhibit 4. Measurement topics included in the NAEP and TIMSS mathematics frameworks: 2003

NAEP ***Measurement topics***	**TIMSS** ***Measurement topics***
Estimate the size of an object or compare objects with respect to given attributes (such as length, area, capacity, volume, weight/mass) Select and use appropriate measurement instruments (for example, manipulatives such as a ruler, meter stick, protractor, thermometer, scales for weight or mass, gauges) Select and use appropriate units of measurement Estimate, calculate (using basic principles or formulas), or compare perimeter, area, volume, and surface area in meaningful contexts to solve mathematical and real-world problems Apply given measurement formulas for perimeter, area, volume, and surface area in problem settings (grades 8 and 12 only) Convert from one measurement to another within the same system (customary or metric) (grades 8 and 12 only) Determine precision, accuracy, and error (grades 8 and 12 only) Make and read scale drawings (grades 8 and 12 only) Select appropriate methods of measurement (such as direct or indirect) Apply the concept of rate to measurement situations (grades 8 and 12 only)	Attributes and units Tools, techniques, and formulas

NOTE: Unless otherwise noted, all topics are intended for all grades (grades 4, 8, and 12 in NAEP and grades 4 and 8 in TIMSS). The number of subtopics or objectives at each grade and level of detail varies across topics and assessments.
SOURCE: U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.*, 2003.

Table 11. Percentage of NAEP and TIMSS fourth- and eighth-grade measurement items classified to the other mathematics assessment framework, by level of content/grade match: 2003

Level of content/grade match	Grade 4		Grade 8	
	NAEP items to TIMSS framework	TIMSS items to NAEP framework	NAEP items to TIMSS framework	TIMSS items to NAEP framework
Total number of measurement items	32	31	30	28
	Percentage distribution			
Classified as same grade	97	81	63	71
Specific match[1] in measurement	66	58	33	61
General match[2] in measurement	19	16	13	7
Match to another content area[3]	13	6	17	4
Classified as another grade[4]	0	16	37	25
Lower grade[5]	†	†	37	25
Higher grade[6]	0	16	†	0
No classification to topics[7]	3	3	0	4

† Not applicable. Grade 4 is the lowest grade in both frameworks; grade 8 is the highest grade in the TIMSS 2003 framework.
[1] Includes items that were classified at the subtopic level at the same grade (and items classified to NAEP framework topics for which no subtopics are included).
[2] Includes items that were classified to a topic but not to a subtopic at the same grade.
[3] Includes items that were classified to a topic or a subtopic in a different content area at the same grade.
[4] Includes items that were classified to a topic or subtopic in any content area at another grade.
[5] Includes grade 8 items classified to grade 4 topics/subtopics.
[6] Includes grade 4 items classified to grade 8 topics/subtopics or grade 8 TIMSS items classified to grade 12 NAEP topics/subtopics.
[7] Includes items that the panel did not classify to a topic at a specific grade level.
NOTE: Data reflect the percentage of items classified by the expert panel at each level. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.*, 2003.

Item distribution across measurement topics

The percentage of *measurement* items across the NAEP and TIMSS *measurement* topics is shown in figures 10 and 11. On the NAEP *measurement* framework, relatively more fourth-grade NAEP items were placed in the topics of *select and use appropriate measurement instruments* (28 percent) and *select and use appropriate units of measurement* (25 percent) than were fourth-grade TIMSS items (10 percent for both topics). A further indication of fourth-grade NAEP items' relative emphasis on selecting and using appropriate units is that on the TIMSS framework, 38 percent of NAEP items (compared to 29 percent in TIMSS) were classified to the topic of *attributes and units*, with a large number of these items classified to the subtopic of *select appropriate standard units to measure length, area, etc*.

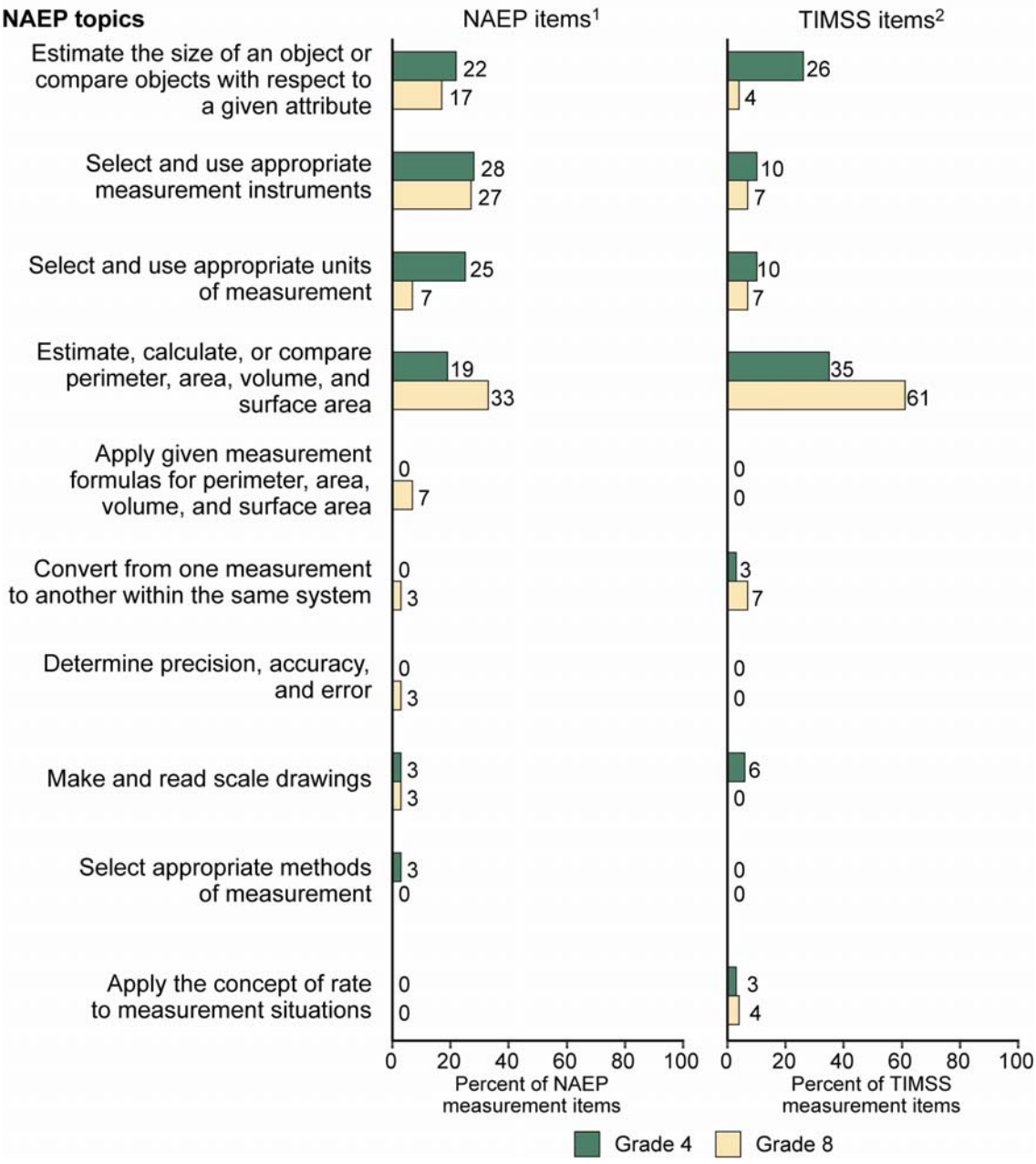
Looking at the two main topics in the TIMSS framework, although more items on both fourth-grade assessments were placed in the topic of *tools, techniques, and formulas* than in *attributes and units*, the emphasis on the former topic is more pronounced in TIMSS than in NAEP. In both assessments, there is still a greater proportion of items in *attributes and units* at the fourth grade than at the eighth grade. Within the *tools, techniques, and formulas* topic, the subtopic with the greatest number of items on both assessments was *compute measurements in simple problem situations*, although there were more TIMSS items addressing this subtopic than NAEP items.

A higher percentage of TIMSS items at both the fourth and eighth grades were classified to the NAEP topic *estimate, calculate, or compare perimeter, area, volume, and surface area,* with 35

percent of TIMSS items compared to 19 percent of NAEP items at fourth grade. At eighth grade, this topic includes the highest proportion of *measurement* items for both NAEP and TIMSS, but the percentage share is much smaller in NAEP (33 percent compared to 61 percent in TIMSS). This potential difference in emphasis is in part supported by the fact that within the fairly broad TIMSS topic of *tools, techniques, and formulas*, close to a third of TIMSS items came from the TIMSS subtopic of *select and use appropriate measurement formulas for perimeter of a rectangle, circumference of a circle, areas of plane figures, surface area and volume of rectangular solid,* compared to very few NAEP items classified by the panel in this subtopic (data not shown). Classification to specific *measurement* subtopics in NAEP indicate that most of the items in this area in both assessments at either grade are related to problems involving properties of two-dimensional shapes rather than three-dimensional objects.

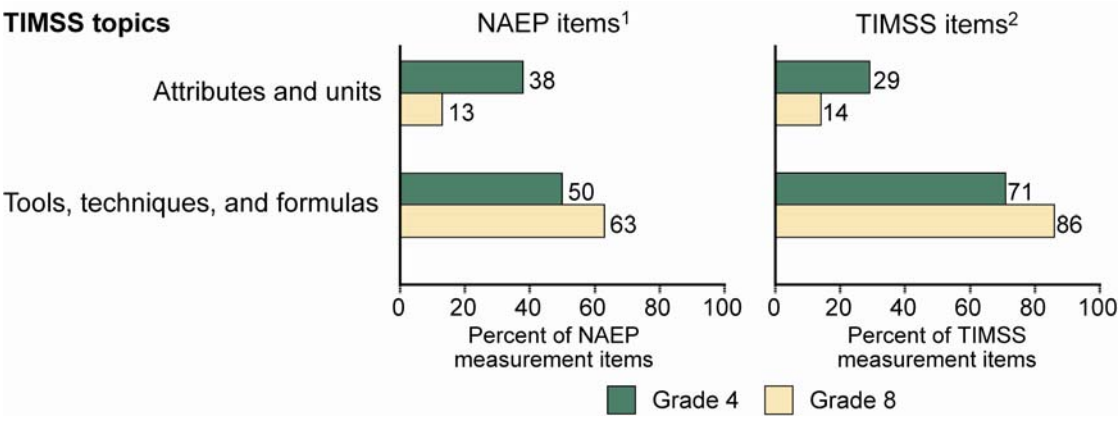
Not surprisingly, there were no items or very few items in either fourth-grade assessment classified to the NAEP topics that were intended for inclusion only at the two higher grades (eighth or twelfth) such as *convert from one measurement to another*, *determine precision, accuracy and error*, and *apply the concept of rate to measurement situations*. Although these NAEP topics were included in the framework at eighth grade, there were also very few eighth-grade items classified to these topics.

Figure 10. Percentage of NAEP and TIMSS measurement items classified to measurement topics in the NAEP mathematics framework, by survey and grade: 2003



[1] NAEP items classified by NAEP developers.
[2] TIMSS items classified by expert panel.
NOTE: Topics may be abbreviated for graphical clarity. Percentages reflect the proportion of *measurement* items classified at either the topic level or the subtopic level at any grade level. Items that were classified to multiple topics were counted in all relevant topics. Bars not shown indicate that no items from that particular grade and assessment were classified to the topic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; and U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002.

Figure 11. Percentage of NAEP and TIMSS measurement items classified to measurement topics in the TIMSS mathematics framework, by survey and grade: 2003



[1] NAEP items classified by expert panel.
[2] TIMSS items classified by TIMSS developers.
NOTE: Percentages reflect the proportion of *measurement* items classified at either the topic level or the subtopic level at any grade level. Items that were classified to multiple topics were counted in all relevant topics.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.*, 2003.

5.3. Geometry

Similar proportions of the fourth- and eighth-grade items in NAEP and TIMSS are in the area of *geometry*. As discussed in section 4, in NAEP, 15 percent of fourth-grade items and 19 percent of eighth-grade items are from the *content strand* of *geometry and spatial sense* (table 4). In TIMSS, 14 percent of fourth-grade items and 17 percent of eighth-grade items are from the *content domain* of *geometry* (table 5). The results in the *geometry* section are based on 28 fourth-grade and 37 eighth-grade items in NAEP, and 21 fourth-grade and 31 eighth-grade items in TIMSS.

Framework comparison in geometry

Although the NAEP and TIMSS *geometry* frameworks are organized somewhat differently, it is difficult to identify clear cases of skills or knowledge addressed in one framework but not the other based on a comparison at the topic level. Exhibit 5 shows a comparison of the *geometry* topics included in the NAEP and TIMSS mathematics frameworks. The NAEP framework consists of nine topics, six of which are further described by subtopics. Only six topics are intended to be addressed at the fourth-grade level. Most NAEP topics address skills and knowledge relevant to a variety of types of geometric figures, with little specification of types of figures (e.g., lines, angles, quadrilaterals). In contrast, of the five *geometry* topics of the TIMSS framework, two relate to specific types of geometric figures (*lines and angles* and *two- and three-dimensional shapes*); a third (*congruence and similarity*) includes subtopics related almost exclusively to triangles.

One potential difference between the frameworks is the relative lack of topics and subtopics in the TIMSS framework that deal with logic and reasoning, ones that would correspond to the NAEP topic of *establish and explain relationships involving geometric concepts* and its related subtopics (*make conjectures*, *validate and justify conclusions and generalizations* and *use informal induction and deduction*). Although the NAEP framework topics and subtopics indicate this difference from TIMSS, it does not appear to exist when the frameworks are implemented in the assessments, as only one NAEP item was classified to this topic by the assessment developers.

Exhibit 5. Geometry topics included in the NAEP and TIMSS mathematics frameworks: 2003

NAEP ***Geometry and spatial sense topics***	**TIMSS** ***Geometry topics***
Describe, visualize, draw, and construct geometric figures	Lines and angles
Investigate and predict results of combining, subdividing, and changing shapes (such as paper folding, dissecting, tilting, rearranging pieces of solids)	Two- and three-dimensional shapes
Identify the relationship (congruence, similarity) between a figure and its image under a transformation	Congruence and similarity
Describe the intersection of two or more geometric figures (grades 8 and 12 only)	Locations and spatial relationships
Classify figures in terms of congruence and similarity, and informally apply these relationships using proportional reasoning where appropriate (grades 8 and 12 only)	Symmetry and transformations
Apply geometric properties and relationships in solving problems	
Establish and explain relationships involving geometric concepts	
Represent problem situations with geometric models and apply properties of figures in meaningful contexts to mathematical and real-world problems	
Represent geometric figures and properties algebraically using coordinates and vectors (grades 8 and 12 only)	

NOTE: Unless otherwise noted, all topics are intended for all grades (grades 4, 8, and 12 in NAEP and grades 4 and 8 in TIMSS). The number of subtopics or objectives at each grade and level of detail varies across topics and assessments.
SOURCE: U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.*, 2003.

Content and grade match in geometry

The degree to which NAEP and TIMSS *geometry* items were matched to topics and subtopics in the other assessment's framework is shown in table 12. At the fourth-grade level, most NAEP and TIMSS *geometry* items were classified to fourth-grade *geometry* topics or subtopics on the other assessment's framework. Although the percentage with either a specific or general match was higher for NAEP than for TIMSS items (82 percent compared to 76 percent), a higher percentage of TIMSS items had a specific match to the NAEP framework than did NAEP items to the TIMSS framework (57 percent compared to 46 percent). Example 15 in appendix E illustrates a fourth-grade TIMSS item that the panel classified in *geometry and spatial sense* but without a good match to a particular NAEP topic. Example 16 in appendix E shows a fourth-grade NAEP item with a general match to the TIMSS topic of *two- and three-dimensional shapes* but not classified to a specific subtopic in the TIMSS *geometry* framework.

At the eighth grade, the level of match of NAEP items to the TIMSS framework was less precise than the match of TIMSS items to the NAEP framework: 57 percent of NAEP items had a general or specific match to eighth-grade TIMSS *geometry* topics or subtopics, compared to 80

percent of TIMSS items. This is the result of a substantial number of eighth-grade NAEP items being classified to fourth-grade TIMSS topics and subtopics, 43 percent. Most of these items were cross-grade items administered at both fourth and eighth grades. In contrast, a number of TIMSS eighth-grade items (13 percent) were classified to topics in the NAEP framework at the twelfth-grade level. Nearly all of the off-grade items were, however, classified to *geometry* topics. A NAEP *geometry* item administered at the fourth, eighth and twelfth grades that was classified at the fourth-grade level on the TIMSS mathematics framework is shown in Example 17. A TIMSS eighth-grade item that was classified as most consistent with the twelfth-grade level of the NAEP framework and specifications document is shown in example 18 in appendix E.

A review of items and their classifications revealed no single reason based on obvious differences between the frameworks that might explain the lack of general or specific match. The items from one assessment that did not match well to the other assessment framework—either because of content or grade classification—came from a variety of original content classifications.

Table 12. Percentage of NAEP and TIMSS fourth- and eighth-grade geometry items classified to the other mathematics assessment framework, by level of content/grade match: 2003

	Grade 4		Grade 8	
Level of content/grade match	NAEP items to TIMSS framework	TIMSS items to NAEP framework	NAEP items to TIMSS framework	TIMSS items to NAEP framework
Total number of geometry items	28	21	37	31
	Percentage distribution			
Classified as same grade	89	81	57	84
Specific match[1] in geometry	46	57	41	45
General match[2] in geometry	36	19	16	35
Match to another content area[3]	7	5	0	3
Classified as another grade[4]	4	10	43	16
Lower grade[5]	†	†	43	3
Higher grade[6]	4	10	†	13
No classification to topics[7]	4	10	0	0

† Not applicable. Grade 4 is the lowest grade in both frameworks; grade 8 is the highest grade in the TIMSS 2003 framework.
[1] Includes items that were classified at the subtopic level at the same grade (and items classified to NAEP framework topics for which no subtopics are included).
[2] Includes items that were classified to a topic but not to a subtopic at the same grade.
[3] Includes items that were classified to a topic or a subtopic in a different content area at the same grade.
[4] Includes items that were classified to a topic or subtopic in any content area at another grade.
[5] Includes grade 8 items classified to grade 4 topics/subtopics.
[6] Includes grade 4 items classified to grade 8 topics/subtopics or grade 8 TIMSS items classified to grade 12 NAEP topics/subtopics.
[7] Includes items that the panel did not classify to a topic at a specific grade level.
NOTE: Data reflect the percentage of items classified by the expert panel at each level. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.*, 2003.

Item distribution across geometry topics

The distribution of items across NAEP and TIMSS *geometry* topics is shown in figures 12 and 13. At the fourth-grade level, on the NAEP framework, the topic in which the highest percentage of both NAEP and TIMSS items were classified was *describe, visualize, draw, and construct geometric figures*. NAEP included relatively more of these items than TIMSS (43 percent compared to 29 percent). Typically, items in this topic required students to use their knowledge of the properties of figures to either answer questions about them or draw geometric figures based on given characteristics.

Another difference between NAEP and TIMSS at the fourth-grade level is that TIMSS includes items related to congruence and similarity whereas NAEP does not. Fourteen percent of TIMSS items were classified to the TIMSS topic of *congruence and similarity*. These same items were classified to the NAEP framework as *classify figures in terms of congruence and apply proportional reasoning*. In contrast, no fourth-grade NAEP items were placed in either the TIMSS topic of *congruence and similarity* or the corresponding NAEP topic which was intended only for assessment at the eighth or twelfth grade.[7]

Using the TIMSS framework as a reference, the distributions of both NAEP and TIMSS eighth-grade items is roughly similar to that for fourth grade. The highest percentage of items on both assessments are classified to the topic of *two- and three-dimensional shapes*, though the relative proportions are greater for the NAEP assessment. Using the NAEP framework, on the other hand, reveals some differences. Although at eighth grade there is still an emphasis with NAEP for items on *describe, visualize, draw, and construct geometric figures,* the percentage of items classified to that topic is lower (22 percent) than at the fourth grade and is closer to the percentage of TIMSS eighth-grade items classified to that topic (16 percent).

Another notable difference is that at the eighth-grade level, a higher percentage of TIMSS items were classified to the NAEP topic of *apply geometric properties and relationships in solving problems*, with more than 40 percent of TIMSS items classified to this topic, compared to 16 percent of NAEP eighth-grade items. In comparison, 5 percent or less of fourth-grade items in either assessment were classified to this topic. Example 19 in appendix E shows an eighth-grade TIMSS item involving the measure of angles in a hexagon that was classified to the topic of *apply geometric properties and relationships in solving problems* in the NAEP mathematics framework.

It should not be concluded, however, that TIMSS places a greater emphasis on problem solving than NAEP, since 11 percent of eighth-grade NAEP items were classified to another NAEP topic that specifically addressed problem solving, *represent problem situations with geometric models and apply properties of figures*. No TIMSS items were placed in this topic.

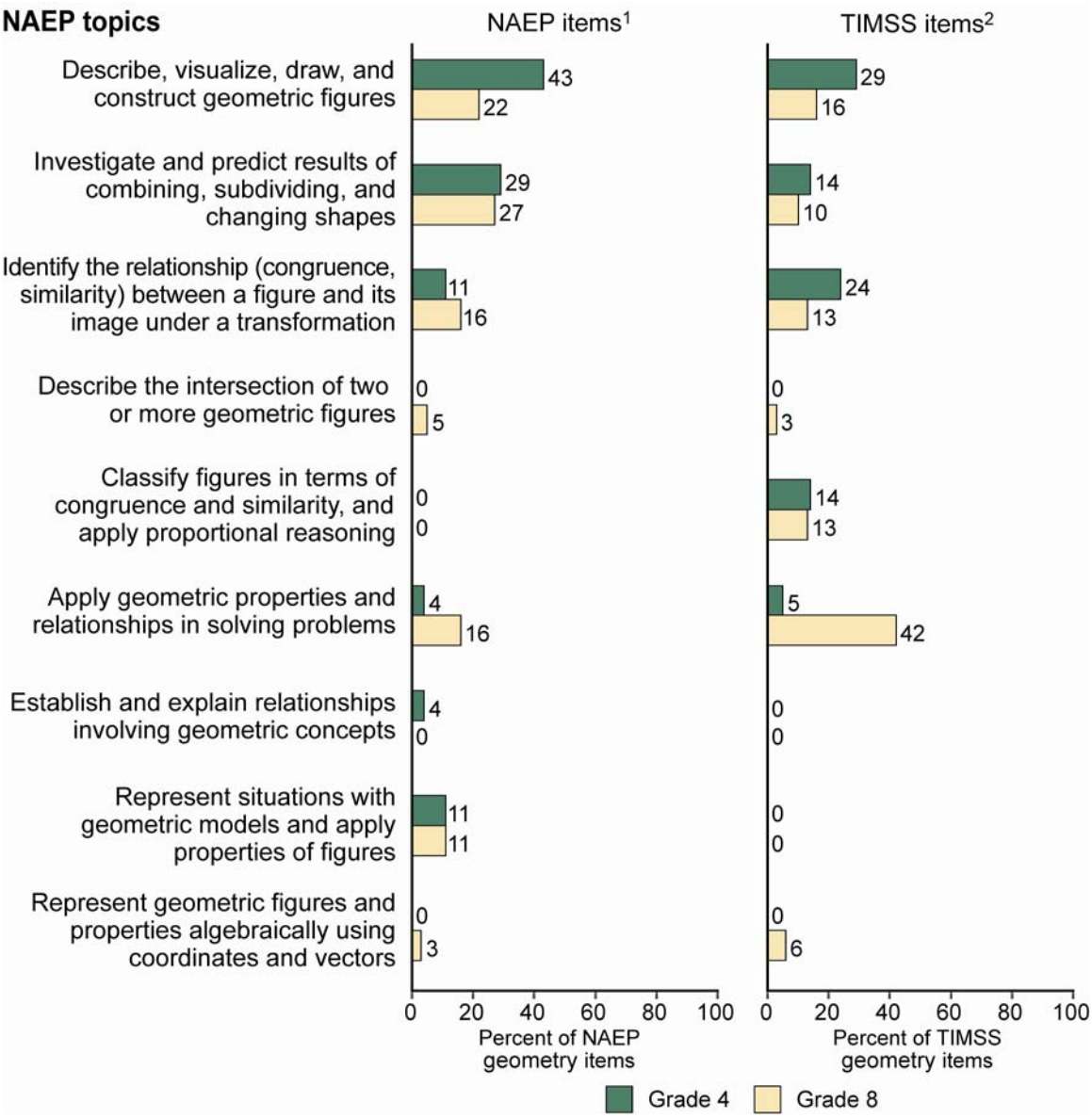
A review of eighth-grade items indicates that the emphasis on problem solving in TIMSS is connected to the inclusion of numerous items that involved finding angle measures, for which the most appropriate topic was one that included the phrase *solving problems*. All but one of the TIMSS items classified to the NAEP framework as *apply geometric properties and relationships in solving problems* required students to find angle measures by relying on, among other things, their

[7] By its name, the NAEP topic of *identify the relationship (congruence, similarity) between a figure and its image under a transformation*, would appear to include items related to congruence and similarity, but in fact none of the NAEP or TIMSS items classified in it dealt with congruence or similarity. Instead, most focused on the transformation aspect of the topic.

knowledge of the properties of angles and geometric figures and algebraic skills (data not shown). In contrast, only one NAEP item involved finding angle measures. This difference does appear to be grounded in the frameworks, as the TIMSS items dealing with angle measures could not be placed in any specific NAEP subtopic.

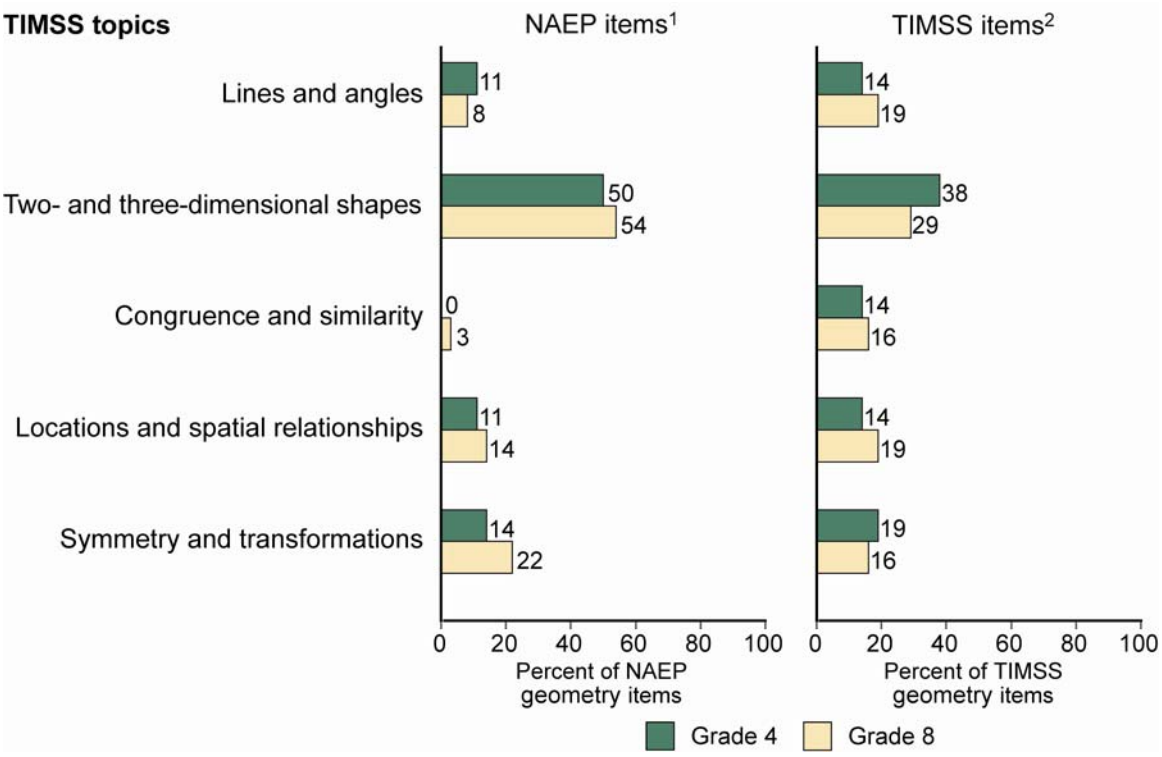
At both the eighth- and the fourth-grade levels, relatively more NAEP items than TIMSS items were classified to the NAEP topic of *investigate and predict results of combining, subdividing, and changing shapes* and to the TIMSS topic of *two- and three-dimensional shapes*. Neither assessment had separate geometry subtopics that facilitated the classification of items on the basis of the use of two- versus three-dimensional shapes. However, further review of the items indicated that there was no substantial difference between the NAEP and TIMSS assessments with respect to their relative emphasis. Both assessments included a relatively small proportion of *geometry* items that involved properties, spatial relationships, or transformations involving three-dimensional objects. This was also true in the *measurement* content area where there were many more items involving measurements of area and perimeter of two-dimensional shapes than volume and surface area of three-dimensional shapes. Both assessments also included some items that required students to recognize the relationship between two-dimensional and three-dimensional shapes (e.g., use of perspective, results of folding, two-dimensional nets of boxes).

Figure 12. Percentage of NAEP and TIMSS geometry items classified to geometry and spatial sense topics in the NAEP mathematics framework, by survey and grade: 2003



[1] NAEP items classified by NAEP developers.
[2] TIMSS items classified by expert panel.
NOTE: Topics may be abbreviated for graphical clarity. Percentages reflect the proportion of *geometry* items classified at either the topic level or the subtopic level at any grade level. Items that were classified to multiple topics were counted in all relevant topics. Bars not shown indicate that no items from that particular grade and assessment were classified to the topic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; and U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002.

Figure 13. Percentage of NAEP and TIMSS geometry items classified to geometry topics in the TIMSS mathematics framework, by survey and grade: 2003



[1] NAEP items classified by expert panel.
[2] TIMSS items classified by TIMSS developers.
NOTE: Percentages reflect the proportion of *geometry* items classified at either the topic level or the subtopic level at any grade level. Items that were classified to multiple topics were counted in all relevant topics. Bars not shown indicate that no items from that particular grade and assessment were classified to the topic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.,* 2003.

5.4. Data

The *data* content area reflects the lowest proportion of both NAEP and TIMSS items at either grade. As discussed in section 4, in NAEP, 10 percent of fourth-grade items and 15 percent of eighth-grade items are from the *content strand* of *data analysis, statistics, and probability* (table 4). In TIMSS, 10 percent of fourth-grade items and 13 percent of eighth-grade items are from the *content domain* of *data* (table 5). The results in the *data* section are based on 19 fourth-grade and 30 eighth-grade items in NAEP, and 15 fourth-grade and 23 eighth-grade items in TIMSS.

Framework comparison in data

Exhibit 6 shows a comparison of the *data* topics included in the NAEP and TIMSS mathematics frameworks. Although the NAEP framework contains more topics than the TIMSS framework (11 compared to 4), in both frameworks there is a clear distinction between topics related to data (e.g., data collection, organization, presentation, and interpretation) and those related to probability. The NAEP framework includes three topics related to probability (*determine the number of ways an event can occur*, *determine the probability of a simple event*, and *apply the basic concept of probability to real-world situations*), while the TIMSS framework includes a single topic of *uncertainty and probability*. The remaining three TIMSS topics cover the *collection*, *organization*, *representation*, and *interpretation of data*. The same could be said of the remaining eight NAEP topics, but it should be noted that three of these topics do not have obvious analogues in the TIMSS framework, either at the topic or subtopic level (*understand and reason about the use and misuse of statistics in our society*, *fit a line or curve to a set of data*, and *design a statistical experiment*).

It is also important to remember that of the numerous NAEP topics and subtopics, only a few are intended to be assessed at the fourth-grade level, and a larger but still limited set is intended for assessment at the eighth-grade level. Two of the topics in the NAEP framework are included at the twelfth grade only—*use measure of central tendency, correlation, dispersion and shapes of distributions to describe statistical relationships* and *fit a line or curve to a set of data and use this line or curve to make predictions about the data, using frequency distributions where appropriate*. In TIMSS, the topic of *uncertainty and probability* is not intended to be addressed at the fourth-grade level.

In the NAEP framework, there appears to be some overlap of subtopics across the topics related to *data*, which could complicate attempts to analyze the distribution across topics. For example, the NAEP topic of *organize and display data and make inferences* includes subtopics and grade-level illustrations that also may correspond to another NAEP topic, *read, interpret, and make predictions using tables and graphs*.

Exhibit 6. Data topics included in the NAEP and TIMSS mathematics frameworks: 2003

NAEP ***Data analysis, statistics, and probability topics***	**TIMSS** ***Data topics***
Read, interpret, and make predictions using tables and graphs	Data collection and organization
Organize and display data and make inferences	Data representation
Understand and apply sampling, randomness, and bias in data collection (grades 8 and 12 only)	Data interpretation
Describe measures of central tendency and dispersion in real-world situations	Uncertainty and probability (grade 8 only)
Use measure of central tendency, correlation, dispersion and shapes of distributions to describe statistical relationships (grade 12 only)	
Understand and reason about the use and misuse of statistics in our society	
Fit a line or curve to a set of data and use this line or curve to make predictions about the data, using frequency distributions where appropriate (grade 12 only)	
Design a statistical experiment to study a problem and communicate the outcomes (grades 8 and 12 only)	
Use basic concepts, trees, and formulas for combinations, permutations, and other counting techniques to determine the number of ways an event can occur (grades 8 and 12 only)	
Determine the probability of a simple event	
Apply the basic concept of probability to real-world situations	

NOTE: Unless otherwise noted, all topics are intended for all grades (grades 4, 8, and 12 in NAEP and grades 4 and 8 in TIMSS). The number of subtopics or objectives at each grade and level of detail varies across topics and assessments.
SOURCE: U.S. Department of Education, National Assessment Governing Board, Mathematics *Framework for the 2003 National Assessment of Educational Progress*, 2002; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.,* 2003.

Content and grade match in data

Table 13 shows the results of the level of content and grade match analyses for the NAEP and TIMSS *data* items. The most striking results from the cross-classification of items to the two frameworks are that 47 percent of fourth-grade NAEP *data* items were classified to topics at the eighth-grade level on the TIMSS framework, whereas all fourth-grade TIMSS items had a specific match to the NAEP fourth-grade framework. Most of the NAEP items that were classified at the eighth-grade level on the TIMSS framework deal with probability, a topic that is not intended to be addressed at the fourth-grade level in TIMSS. An additional 11 percent of NAEP items were not given a grade classification at the topic level on the TIMSS framework.

Of the fourth-grade NAEP items that were placed at the fourth-grade level on the TIMSS framework, most had a specific match (37 percent of all items, compared to 5 percent that had a general match).

Similar to the fourth-grade items, eighth-grade TIMSS *data* items were classified to the NAEP framework more precisely than NAEP items to the TIMSS framework, although the contrast was not as great: for TIMSS, more than 80 percent of items had a specific match to the NAEP framework compared to the 67 percent of NAEP items with a specific match to the TIMSS framework.

Unlike the fourth-grade assessment, where the high percentage of NAEP items dealing with probability resulted in a relatively high percentage of items classified to a higher grade level, at the eighth-grade level, there is no single reason for the lack of specific match between NAEP items and the TIMSS framework. It is worth noting, however, that some of the NAEP items that were not given a subtopic classification on the TIMSS framework required students to compute either the median or the mode of a single dataset. The related TIMSS subtopics were strictly interpreted by the panel to be limited to comparisons of measures of central tendency across data sets. Other NAEP items were found to deal with combinations, a common area in probability, but not one mentioned explicitly in the TIMSS subtopics in *uncertainty and probability*.

Table 13. Percentage of NAEP and TIMSS fourth- and eighth-grade data items classified to the other mathematics assessment framework, by level of content/grade match: 2003

Level of content/grade match	Grade 4		Grade 8	
	NAEP items to TIMSS framework	TIMSS items to NAEP framework	NAEP items to TIMSS framework	TIMSS items to NAEP framework
Total number of data items	19	15	30	23
	Percentage distribution			
Classified as same grade	42	100	83	91
Specific match[1] in data	37	100	67	83
General match[2] in data	5	0	17	0
Match to another content area[3]	0	0	0	9
Classified as another grade[4]	47	0	10	4
Lower grade[5]	†	†	10	4
Higher grade[6]	47	0	†	0
No classification to topics[7]	11	0	7	4

† Not applicable. Grade 4 is the lowest grade in both frameworks; grade 8 is the highest grade in the TIMSS 2003 framework.
[1] Includes items that were classified at the subtopic level at the same grade (and items classified to NAEP framework topics for which no subtopics are included).
[2] Includes items that were classified to a topic but not to a subtopic at the same grade.
[3] Includes items that were classified to a topic or a subtopic in a different content area at the same grade.
[4] Includes items that were classified to a topic or subtopic in any content area at another grade.
[5] Includes grade 8 items classified to grade 4 topics/subtopics.
[6] Includes grade 4 items classified to grade 8 topics/subtopics or grade 8 TIMSS items classified to grade 12 NAEP topics/subtopics.
[7] Includes items that the panel did not classify to a topic at a specific grade level.
NOTE: Data reflect the percentage of items classified by the expert panel at each level. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.,* 2003.

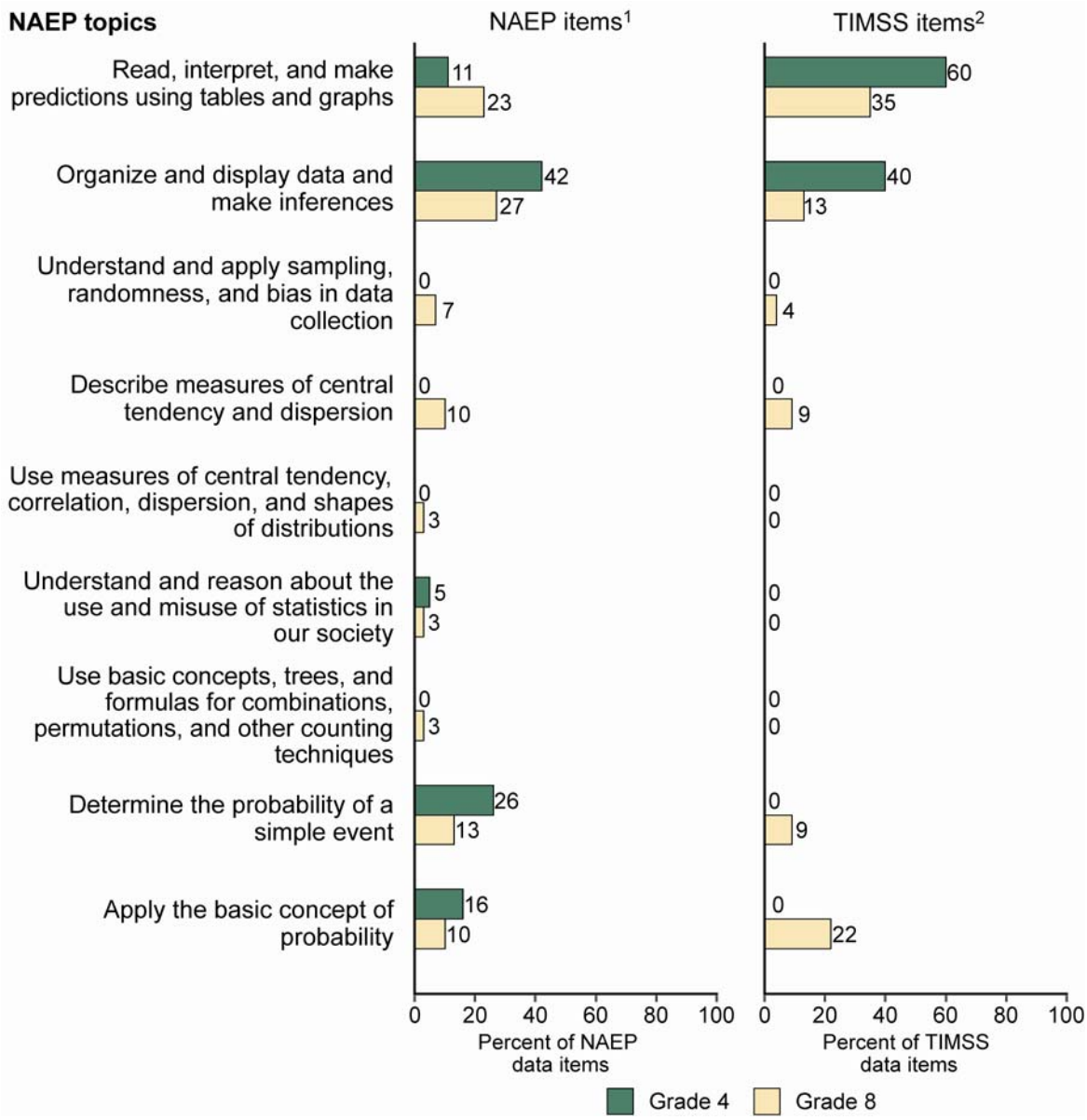
Item distribution across data topics

Figures 14 and 15 show the distribution of NAEP and TIMSS *data* items across the *data* topics included in the two frameworks. As noted above, the most notable difference between NAEP and TIMSS fourth-grade *data* items is that NAEP includes several items dealing with probability (about 40 percent), whereas TIMSS does not include any at the fourth grade. These NAEP items are clearly identifiable both in their classifications to the NAEP framework (in the topics of *determine the probability of a simple event* and *apply the basic concept of probability to real-world situations*) and in their cross-classification to the TIMSS framework, in the topic of *uncertainty and probability*.

More fourth-grade TIMSS items were classified to the NAEP topic of *read, interpret, and make predictions using tables and graphs* than were NAEP items (60 percent compared to 11 percent). Within this topic, both NAEP and TIMSS include a few items that require students to perform some sort of computation based on information provided in a table or graph. However, most of the TIMSS items classified to this topic required students to either answer a straightforward question based on a given graph or table or choose from a set of tables or graphs the one that best represents a set of data. Most of these items were originally classified to the TIMSS framework in the topic of *data representation*, which helps explain the higher percentage of TIMSS items classified to that TIMSS topic. A TIMSS *data* item of this nature is demonstrated by Example 20 in appendix E.

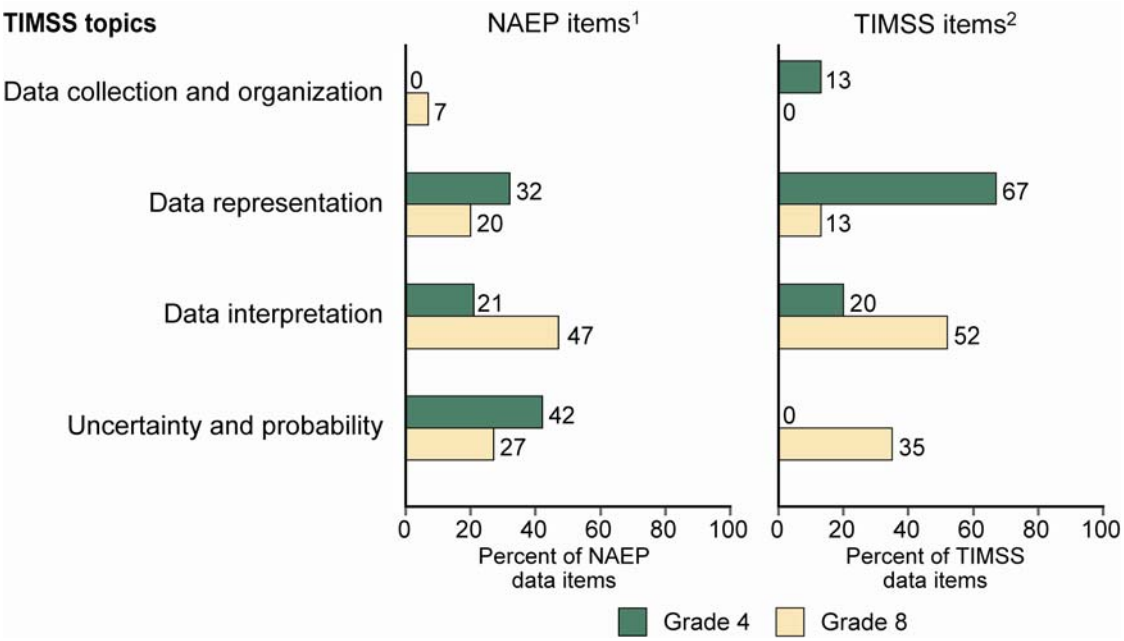
For NAEP and especially TIMSS, the relative emphasis on *data representation* at the fourth grade shifts to an increased emphasis on *data interpretation* at the eighth grade. Compared to fourth grade, at the eighth-grade, there were less dramatic differences between the distribution of NAEP and TIMSS *data* items across both the NAEP and TIMSS frameworks, with a large concentration of items in topics related to the display or interpretation of data. At the eighth grade, NAEP and TIMSS had more comparable numbers of items covering topics related to probability. One of the differences between the two eighth-grade assessments is the higher percentage of NAEP items classified to the NAEP topic of *organize and display data and make inferences* (27 percent compared to 13 percent in TIMSS).

Figure 14. Percentage of NAEP and TIMSS data items classified to data analysis, statistics, and probability topics in the NAEP mathematics framework, by survey and grade: 2003



[1] NAEP items classified by NAEP developers.
[2] TIMSS items classified by expert panel.
NOTE: Topics may be abbreviated for graphical clarity. Two NAEP framework topics included for assessment at eighth and/or twelfth grade(s) only are not reflected in this figure, as no grade 4 or grade 8 items in either assessment were classified to these topics: *fit a line or curve to a set of data and use this linear curve to make predictions about the data, using frequency distributions where appropriate* and *design a statistical experiment to study a problem and communicate the outcomes*. Percentages reflect the proportion of *data* items classified at either the topic level or the subtopic level at any grade level. Items that were classified to multiple topics were counted in all relevant topics. Bars not shown indicate that no items from that particular grade and assessment were classified to the topic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; and U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002.

Figure 15. Percentage of NAEP and TIMSS data items classified to data topics in the TIMSS mathematics framework, by survey and grade: 2003



[1] NAEP items classified by expert panel.
[2] TIMSS items classified by TIMSS developers.
NOTE: Percentages reflect the proportion of *data* items classified at either the topic level or the subtopic level at any grade level. Items that were classified to multiple topics were counted in all relevant topics. Bars not shown indicate that no items from that particular grade and assessment were classified to the topic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.,* 2003.

5.5. Algebra

Algebra items make up a larger proportion of the eighth-grade than the fourth-grade items for both NAEP and TIMSS. As discussed in section 4, in NAEP, 14 percent of fourth-grade items and 25 percent of eighth-grade items are from the *content strand* of *algebra and functions* (table 4). In TIMSS, 16 percent of fourth-grade items and 23 percent of eighth-grade items are from the *content domain* of *algebra* (table 5). The results in the *algebra* section are based on 26 fourth-grade and 49 eighth-grade items in NAEP, and 23 fourth-grade and 42 eighth-grade items in TIMSS.

Framework comparison in algebra

The NAEP and TIMSS *algebra* frameworks are compared in exhibit 7. Exhibit 7 shows a comparison of the *algebra* topics included in the NAEP and TIMSS mathematics frameworks. The NAEP framework includes several topics and subtopics included exclusively or primarily for the twelfth-grade assessment that are not addressed in the TIMSS framework. Among these are the topics of *solve polynomial equations with real and complex roots using a variety of algebraic and graphical methods and using appropriate tools*, *apply function concepts to model and deal with real-world situations*, and *use trigonometry*. As in the other content areas, the TIMSS framework includes a few broad topics with a number of grade-specific subtopics within each. Only the *algebraic expressions* topic is not to be assessed at the fourth grade in TIMSS.

Considering only the NAEP topics and subtopics included for the fourth and eighth grades, there is rough agreement between the two frameworks regarding *algebra* content, with a few exceptions. The topics related to patterns appear to be covered in both frameworks, in TIMSS under the topic of *patterns* and in NAEP under the topic of *describe, extend, interpolate, transform, and create a wide variety of patterns and functional relationships*. However, although functions and functional thinking are addressed by both frameworks, the correspondence between NAEP and TIMSS topics and subtopics is not obvious.

There is also some correspondence between the TIMSS topics of *algebraic expressions* and *equations and formulas* and the NAEP topics of *represent and describe solutions to linear equations and inequalities* and *interpret contextual situations and perform algebraic operations on real numbers and algebraic expressions to solve mathematical and real-world problems*. In TIMSS, however, the subtopics indicate a focus on simplifying and evaluating algebraic expressions whereas the subtopics in the NAEP framework indicate a greater focus on problem solving.

One clear difference is that the NAEP framework includes an explicit *algebra* topic and related subtopics (*make conjectures, validate and justify conclusions and generalizations, use formal induction, and deduction*) that deal with *mathematical reasoning* whereas the TIMSS *algebra content domain* does not. Rather, the TIMSS framework specifies related abilities within its *cognitive domain of reasoning* that goes across all *content domains*. Another difference is that the NAEP subtopic of *identify or graph sets of points on a number line or in a rectangular coordinate system* appears more similar to a subtopic of the TIMSS *geometry* topic of *locations and spatial relationships* (*locate points using number lines, coordinate grids, maps*) than any TIMSS topic or subtopic found in *algebra*.

Exhibit 7. Algebra topics included in the NAEP and TIMSS mathematics frameworks: 2003

NAEP ***Algebra and functions topics***	**TIMSS** ***Algebra topics***
Describe, extend, interpolate, transform, and create a wide variety of patterns and functional relationships	Patterns
Use multiple representations for situations to translate among diagrams, models, and symbolic expressions	Algebraic expressions (grade 8 only)
Use number lines and rectangular coordinate systems as representational tools	Relationships
Represent and describe solutions to linear equations and inequalities to solve mathematical and real-world problems	Equations and formulas
Interpret contextual situations and perform algebraic operations on real numbers and algebraic expressions to solve mathematical and real-world problems (grades 8 and 12 only)	
Solve systems of equations and inequalities using appropriate methods (grades 8 and 12 only)	
Use mathematical reasoning	
Represent problem situations with discrete structures (grades 8 and 12 only)	
Solve polynomial equations with real and complex roots using a variety of algebraic and graphical methods using appropriate tools (grade 12 only)	
Approximate solutions of equations (bisection, sign changes, successive approximations) (grades 8 and 12 only)	
Use appropriate notation and terminology to describe functions and their properties (grade 12 only)	
Compare and apply the numerical, symbolic, and graphical properties of a variety of functions and families of functions, examining general parameters and their effect on curve shape (grades 8 and 12 only)	
Apply function concepts to model and deal with real-world situations (grades 8 and 12 only)	
Use trigonometry (grade 12 only)	

NOTE: Unless otherwise noted, all topics are intended for all grades (grades 4, 8, and 12 in NAEP and grades 4 and 8 in TIMSS). The number of subtopics or objectives at each grade and level of detail varies across topics and assessments.
SOURCE: U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.,* 2003.

Content and grade match in algebra

As seen in table 14, nearly all fourth-grade level NAEP and TIMSS *algebra* items were classified to fourth-grade topics in the other assessment framework (96 percent for both). Of these items, the majority had a specific match to the other assessment's framework, although the percentage of NAEP items with a specific match to the TIMSS framework was higher than the

percentage of TIMSS items with a specific match to the NAEP framework (81 percent compared to 65 percent).

At the eighth-grade level, the classification of NAEP items to the TIMSS framework was less precise than that for the TIMSS items to the NAEP framework. While 95 percent of eighth-grade TIMSS items were classified at the eighth-grade level on the NAEP framework (93 percent with a specific match to *algebra* subtopics), 78 percent of eighth-grade NAEP items were classified to eighth-grade topics in TIMSS (57 percent had a specific match to algebra subtopics). The lower level of match between NAEP items and the TIMSS framework is due in large measure to the 20 percent of items classified in *content domains* other than *algebra*. This indicates that the NAEP framework and NAEP items reflect a broader conception of *algebra* than the TIMSS framework or items, including topics that in TIMSS would be found in other *content domains*. Example 21 in appendix E shows an eighth-grade NAEP *algebra* item that was classified to the *number content domain* in TIMSS. This item was from the NAEP *algebra* subtopic of *perform basic operations using appropriate tools, on real numbers in meaningful contexts (including grouping and order of multiple operations involving basic operations, exponents, and roots)*. In addition, nearly 20 percent of NAEP eighth-grade items were classified to fourth-grade TIMSS topics and subtopics. Most were classified to either of the two *algebra* topics of *patterns* and *equations and formulas*. Five percent of TIMSS eighth-grade *algebra* items were classified at the twelfth-grade level according to the NAEP framework.

Table 14. Percentage of NAEP and TIMSS fourth- and eighth-grade algebra items classified to the other mathematics assessment framework, by level of content/grade match: 2003

Level of content/grade match	Grade 4		Grade 8	
	NAEP items to TIMSS framework	TIMSS items to NAEP framework	NAEP items to TIMSS framework	TIMSS items to NAEP framework
Total number of algebra items	26	23	49	42
	Percentage distribution			
Classified as same grade	96	96	78	95
Specific match[1] in algebra	81	65	57	93
General match[2] in algebra	0	17	0	0
Match to another content area[3]	15	13	20	2
Classified as another grade[4]	0	4	18	5
Lower grade[5]	†	†	18	0
Higher grade[6]	0	4	†	5
No classification to topics[7]	4	0	4	0

† Not applicable. Grade 4 is the lowest grade in both frameworks; grade 8 is the highest grade in the TIMSS 2003 framework.
[1] Includes items that were classified at the subtopic level at the same grade (and items classified to NAEP framework topics for which no subtopics are included).
[2] Includes items that were classified to a topic but not to a subtopic at the same grade.
[3] Includes items that were classified to a topic or a subtopic in a different content area at the same grade.
[4] Includes items that were classified to a topic or subtopic in any content area at another grade.
[5] Includes grade 8 items classified to grade 4 topics/subtopics.
[6] Includes grade 4 items classified to grade 8 topics/subtopics or grade 8 TIMSS items classified to grade 12 NAEP topics/subtopics.
[7] Includes items that the panel did not classify to a topic at a specific grade level.
NOTE: Data reflect the percentage of items classified by the expert panel at each level. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.,* 2003.

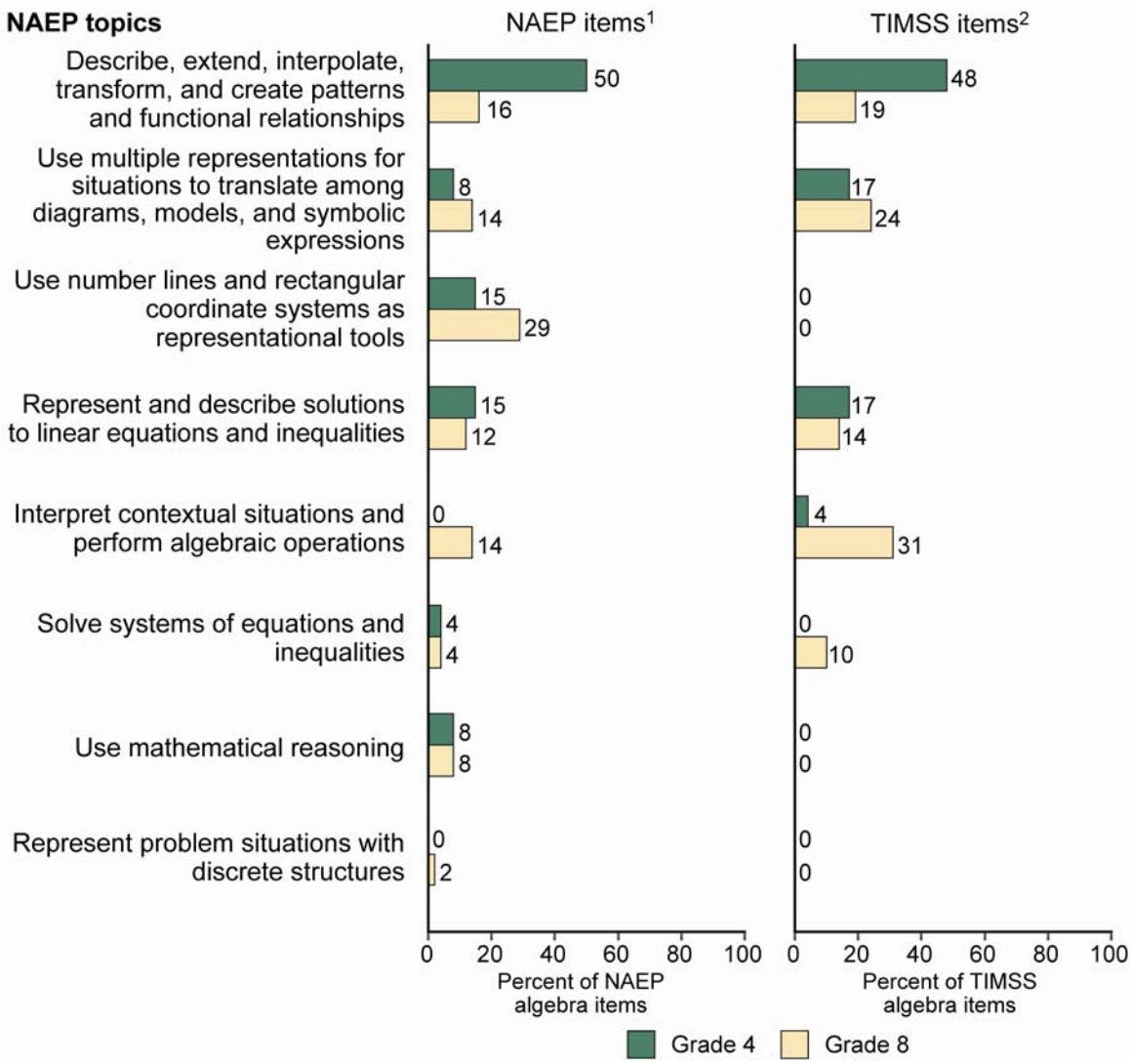
Item distribution across algebra topics

Figures 16 and 17 show the distribution of *algebra* items from each assessment across *algebra* topics in the NAEP and TIMSS framework, respectively. At the fourth-grade level, NAEP and TIMSS items had somewhat similar distributions across NAEP topics, with approximately half of the items on both assessments classified to the topic of *describe, extend, interpolate, transform, and create patterns and functional relationships*. This emphasis on patterns is reflected in the TIMSS framework as well, with 39 percent of items on both assessments classified to the topic of *patterns* and an additional 17 percent of TIMSS items classified to the topic of *relationships*. No fourth-grade NAEP items were classified to the TIMSS topic of *relationships*. Although it is possible that some NAEP items did in fact address this topic but were classified elsewhere, the fact that the panel did place NAEP items in this topic at the eighth-grade level indicates that they would have done so at the fourth-grade level as well had they found any items that matched this TIMSS topic. Neither NAEP nor TIMSS had any fourth-grade items classified to the TIMSS topic of *algebraic expressions*, which is consistent with the definition in the TIMSS framework of this as a topic appropriate for the eighth grade only.

One notable difference at both the fourth- and eighth-grade levels is that sizeable percentages of NAEP items (15 percent at the fourth-grade level and 29 percent at the eighth-grade level) were classified to the NAEP topic of *use number lines and rectangular coordinate systems as representational tools*. In contrast, no TIMSS *algebra* items at either grade level were placed in this topic. This difference at the item level reflects one of the differences in the *algebra* frameworks discussed above, that this NAEP topic is perhaps more closely aligned with TIMSS topics and subtopics in the *geometry content domain*. In fact, when the panel placed these NAEP items on the TIMSS framework, they placed a number of the fourth-grade items in TIMSS *geometry* topics and some of the eighth-grade items in a single subtopic in the TIMSS *geometry* framework (*locate points using number lines, coordinate grids, maps*), which is reflected in the 15–20 percent of NAEP *algebra* items matched to another content area in table 11.

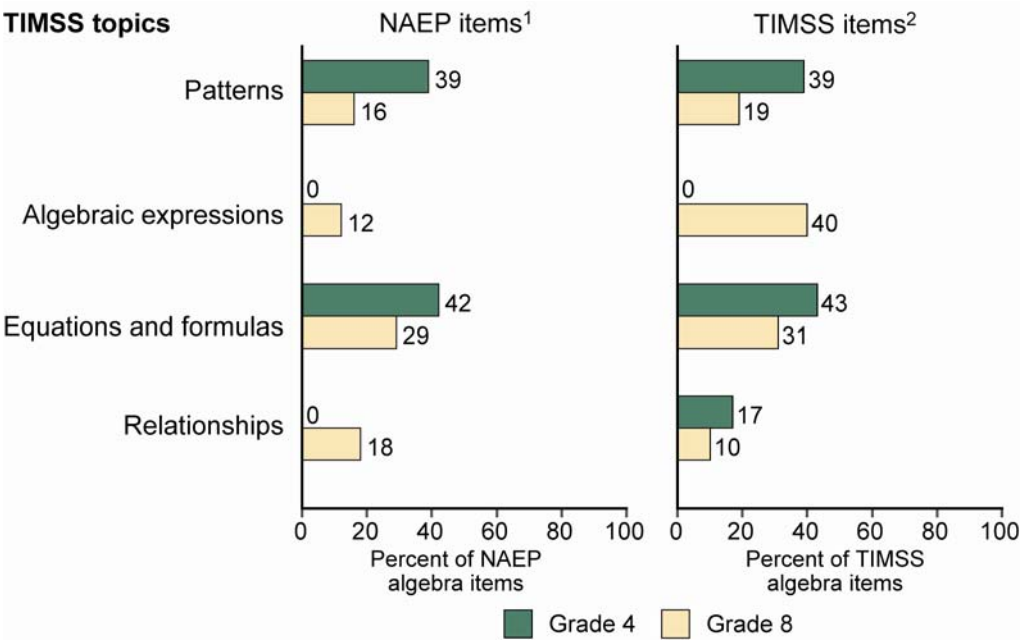
Another difference at the eighth-grade level is that 40 percent of TIMSS items were classified to the TIMSS topic of *algebraic expressions*, compared to 12 percent of NAEP items. All but one of these TIMSS items were classified to the NAEP topics of either *use multiple representations for situations to translate among diagrams, models, and symbolic expressions* or *interpret contextual situations and perform algebraic operations*, which might explain the relatively higher percentages of TIMSS items placed in those NAEP topics.

Figure 16. Percentage of NAEP and TIMSS algebra items classified to algebra and functions topics in the NAEP mathematics framework, by survey and grade: 2003



[1] NAEP items classified by NAEP developers.
[2] TIMSS items classified by expert panel.
NOTE: Topics may be abbreviated for graphical clarity. Six NAEP framework topics included for assessment at eighth and/or twelfth grade(s) only are not reflected in this figure, as no grade 4 or grade 8 items in either assessment were classified to these topics: *solve polynomial equations with real and complex roots using a variety of algebraic and graphical methods and using appropriate tools*; *approximate solutions of equation (bisection, sign changes, successive approximations)s*; *use appropriate notation and terminology to describe functions and their properties*; *compare and apply the numerical, symbolic, and graphical properties of a variety of functions and families of functions, examining general parameters and their effect on curve shape*; *apply function concepts to model and deal with real-world situations*; and *use trigonometry*. Percentages reflect the proportion of *algebra* items classified at either the topic level or the subtopic level at any grade level. Items that were classified to multiple topics were counted in all relevant topics. Bars not shown indicate that no items from that particular grade and assessment were classified to the topic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; and U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002.

Figure 17. Percentage of NAEP and TIMSS algebra items classified to algebra topics in the TIMSS mathematics framework, by survey and grade: 2003



[1] NAEP items classified by expert panel.
[2] TIMSS items classified by TIMSS developers.
NOTE: Percentages reflect the proportion of *algebra* items classified at either the topic level or the subtopic level at any grade level. Items that were classified to multiple topics were counted in all relevant topics. Bars not shown indicate that no items from that particular grade and assessment were classified to the topic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; International Association for the Evaluation of Educational Achievement, Trends in International Mathematics and Science Study (TIMSS) 2003 Assessment; and International Study Center, Lynch School of Education, Boston College, *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed.,* 2003.

This section compared the content of the NAEP and TIMSS assessments in each of the main content areas of *number*, *measurement*, *geometry*, *data*, and *algebra*. The next section compares PISA items with NAEP *problem solving* items from the fourth- and eighth-grade assessments. Comparisons are made with respect to content coverage, *competency cluster* and *situations or contexts* dimensions from the PISA framework, and general item characteristics.

6. NAEP/PISA Comparisons

As described in the methods section, a subset of the panel conducted a comparison of PISA items and a specially selected set of NAEP items. The selected set of NAEP items were drawn from the 2003 eighth-grade and the 2000 twelfth-grade NAEP assessments.[8] They included all items from the *mathematical abilities* category of *problem solving* and all items requiring an extended response. All of the extended-response items were from the NAEP *problem solving* category except one item that was classified as *conceptual understanding*. The NAEP and PISA items were compared on several dimensions of the other assessment's framework. Some of the results relating to PISA items are presented in the earlier section on overall comparisons (section 4). PISA items were not included in the sections devoted to each of the content areas (section 5) because PISA is not designed to correspond as closely to the curriculum-based content areas defined by NAEP and TIMSS. This section presents some additional comparisons between NAEP and PISA items related to mathematical content covered in terms of both the NAEP *content strands* and the PISA *overarching ideas* as well as comparisons based on the *competency clusters* and *situations or contexts* dimensions defined in the PISA framework.[9]

It is important to emphasize that the NAEP items included in these comparisons are not a complete set of items for either eighth or twelfth grade, but rather are selected items across the two grade levels that are closest in age to the PISA target population (15-year-olds). Therefore, comparisons made between NAEP and PISA items in this section should not be interpreted as representative of NAEP overall. Instead, the purpose of these comparisons is to compare PISA items with NAEP items that were designed to measure students' problem-solving abilities. Unlike the previous sections, this section includes comparisons based on the dimensions in the PISA framework, which, because of its emphasis on the application of knowledge and skills to real-life problem solving, may provide an additional perspective for examining the NAEP items.

6.1. Content Comparisons Based on NAEP and PISA Frameworks

Although the correspondence between the content dimensions defined in the NAEP and PISA frameworks is not as strong as between NAEP and TIMSS, there was nevertheless considerable overlap of the items from each assessment with the other assessment framework from a strictly mathematics content perspective. In fact, as discussed in the section on overall comparisons (tables 4 and 7), all of the PISA items were classified to at least the broad *content strand* level, with 92 percent classified to a specific subtopic in the NAEP framework. As displayed in tables 15 and 16, there is a strong correspondence between the PISA *overarching idea* of *uncertainty* and the NAEP *content strand* of *data analysis, statistics, and probability*. Ninety-five percent of PISA *uncertainty* items and 81 percent of NAEP *data analysis, statistics, and probability* items were classified to the corresponding content area in the other framework. There also appears to be a strong correspondence between *space and shape* in PISA and *measurement* and *geometry and spatial sense* in NAEP. Most PISA items from *space and shape* (90 percent) and NAEP items from *geometry and spatial sense* (95 percent) and *measurement* (67 percent) were cross-classified to these corresponding content areas across NAEP and PISA frameworks. Even for the *overarching idea*s of *change and relationships* and *quantity*, which have less direct correspondence with a particular NAEP *content strand*, PISA items in these areas still mapped to NAEP topics across the *content strands*.

[8] The 2000 assessment was the most recent twelfth-grade NAEP assessment at the time of the expert panel meeting.
[9] For all comparisons based on the PISA framework dimensions, NAEP item classifications were made by the expert panel, while those for PISA items reflect classification information provided by the PISA assessment developers.

Additional information about the distribution of PISA items across NAEP mathematics topics is provided in the supplementary data in appendix D (table D-3). While the *uncertainty* items in PISA are classified across six different topics in the NAEP *content strand* of *data analysis, statistics, and probability*, the majority of items relate to three topics: *read, interpret, and make predictions using tables and graphs*; *describe measures of central tendency and dispersion*; and *understand and reason about the use and misuse of statistics in our society*. The PISA items in *space and shape* cover a range of *measurement* and *geometry* topics in NAEP, with particular focus on the *measurement* topic of *estimate, calculate or compare perimeter, area, volume, and surface area.* Items in *change and relationships* have the greatest focus on the topic *read, interpret, and make predictions using tables and graphs.* The *quantity* items cover a couple of topics in each of the NAEP *content strands* except *geometry*. Not surprisingly, a number of these PISA items were classified as *use computations and estimation in applications*.

In general, the NAEP items were more difficult to match to the PISA framework than vice-versa. While all PISA items were classified to a NAEP *content strand* and most to a specific NAEP subtopic, a number of NAEP items could not be placed in any PISA *overarching idea* (table 16). These items came from three NAEP *content strands*: *number sense, properties, and operations* (12 percent); *geometry and spatial sense* (5 percent), and *algebra and functions* (11 percent). Two of these items are related to logic, a topic not found on the PISA framework, and the other items covered a range of topics in NAEP. There also were NAEP items in all content areas except *geometry* that were classified to multiple *overarching ideas* in PISA.

Table 15. Percentage distribution of PISA mathematics items across NAEP mathematics content strands, by PISA overarching idea: 2003

NAEP content strand	Total	PISA overarching idea			
		Change and relationships	Quantity	Space and shape	Uncertainty
Number of PISA items	85	22	23	20	20
		Percentage distribution			
Number sense, properties, and operations	22	14	57	10	5
Measurement	16	9	13	45	0
Geometry and spatial sense	11	0	0	45	0
Data analysis, statistics, and probability	39	41	22	0	95
Algebra and functions	9	27	9	0	0
Classified to multiple strands	2	9	0	0	0

NOTE: Percentages reflect the proportion of PISA items classified by the expert panel to the NAEP 2003 mathematics *content strands* at any level of specificity (*content strand*, *topic*, or *subtopic*) and at any grade level. Of the items classified to multiple strands, one item was classified to both *geometry and spatial sense* and to *algebra and functions*, and one was classified to *measurement* and to *data analysis, statistics, and probability*. Detail may not sum to totals because of rounding.
SOURCE: Organization for Economic Cooperation and Development, Program for International Student Assessment (PISA) 2003 Mathematical Literacy Assessment; U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002; and Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills*, 2003.

Table 16. Percentage distribution of NAEP 2003 eighth-grade and NAEP 2000 twelfth-grade mathematics problem solving items across PISA overarching ideas, by NAEP content strand: 2000 and 2003

PISA overarching idea	Total	NAEP content strand: Number sense, properties, and operations	Measurement	Geometry and spatial sense	Data analysis, statistics, and probability	Algebra and functions
Number of NAEP problem solving items[1]	79	17	9	19	16	18
			Percentage distribution			
Change and relationships	28	47	0	0	13	67
Quantity	9	35	11	0	0	0
Space and shape	32	0	67	95	0	11
Uncertainty	16	0	0	0	81	0
Classified to multiple overarching ideas	8	6	22	0	6	11
Not classified to a PISA overarching idea	6	12	0	5	0	11

[1] NAEP data are based on a selected set of 79 mathematics items from the NAEP 2003 grade 8 and NAEP 2000 grade 12 assessments. This selected set of items consists of all extended-response items and all items from the *problem solving* category of the NAEP 2003 framework. All extended-response items are from the *problem solving* category except one item from the *conceptual understanding* category.
NOTE: Percentages reflect the proportion of NAEP items classified by the expert panel to each *overarching idea* category in the PISA 2003 mathematical literacy framework. Of the items classified to multiple *overarching ideas*, five items were classified to both *change and relationships* and *quantity*, and one item was classified to both *space and shape* and *quantity*. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2000 Mathematics Assessment; U.S. Department of Education, National Assessment Governing Board, *Mathematics Framework for the 2003 National Assessment of Educational Progress*, 2002; and Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills,* 2003.

6.2. PISA Competency Clusters

As described in the overview section (section 2), the PISA framework defines three *competency clusters*. The *reproduction* cluster involves the reproduction of practiced knowledge, including the recall of facts, execution of routine procedures, and use of standard solution and thinking strategies. The *connections* cluster involves familiar settings similar to the *reproduction* cluster, but goes beyond it by involving solution strategies that are not routine. The *reflection* cluster requires more planning and original thought than do the other two clusters.

The fact that the NAEP items that were compared to the PISA framework were those from the *problem solving* category does not necessarily mean that all would be classified to a particular PISA cluster. All three clusters can involve problem solving and are distinguished more by the balance between the demand for recall of facts or procedures versus more creative thinking and solution strategies. In fact, only 4 percent of the NAEP *problem solving* items overall were classified to PISA's *reflection* cluster, compared to 22 percent of PISA items (table 17). The remaining NAEP items were split between the *reproduction* and the *connections* clusters. When considering the overall set of the NAEP *problem solving* items for both eighth and twelfth grades, a similar proportion of NAEP and PISA items are in the *connections* category (44 and 47 percent), while a greater proportion of NAEP items were classified as *reproduction* (46 percent compared to 31 percent for PISA). A breakdown by grade level for the NAEP items reveals a greater percentage in

eighth-grade classified to the *reproduction* cluster, and less to the *connections* cluster compared to twelfth grade. Five percent or less of items from both grades was classified to the *reflection* cluster. Examples 22 and 23 in appendix E illustrate NAEP *problem solving* items classified to the PISA *competency clusters* of *reproduction* and *connections*, respectively. Both of these items are cross-grade items administered at both the eighth and twelfth grades.

Table 17. Percentage distribution of PISA 2003 mathematics items and NAEP 2003 eighth-grade and NAEP 2000 twelfth-grade mathematics problem solving items classified to PISA competency clusters, by grade/age: 2000 and 2003

	PISA[1]	NAEP[2]		
PISA competency cluster	15-year-olds	Total[3]	Grade 8	Grade 12
Reproduction	31	46	58	39
Connections	47	44	36	48
Reflection	22	4	2	5

[1] PISA items classified by PISA developers.
[2] NAEP items classified by expert panel.
[3] NAEP eighth- and twelfth-grade items combined.
NOTE: NAEP data are based on a selected set of 79 mathematics items from the NAEP 2003 grade 8 and NAEP 2000 grade 12 assessments. This selected set of items consists of all extended-response items and all items from the *problem solving* category of the NAEP 2003 framework. All extended-response items are from the *problem solving* category except one item from the *conceptual understanding* category. Percentages reflect the proportion of items classified to each *competency cluster* in the PISA 2003 mathematical literacy framework. Five NAEP items that the panel did not classify to the PISA framework are not included. NAEP cross-grade items administered at grade 8 and grade 12 are reflected in the percentages for both grades, but only counted once in the total column. Detail may not sum to totals because of rounding or omitted items.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2000 Mathematics Assessment; Organization for Economic Cooperation and Development, Program for International Student Assessment (PISA) 2003 Mathematical Literacy Assessment; and Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills*, 2003.

Table 18 displays the distribution of NAEP *problem solving* items across PISA *competency clusters* by NAEP *content strand*. These results show that all of the NAEP items classified to the *reflection* cluster came from the *content strands* of *geometry and spatial sense* and *data analysis, statistics, and probability*. In addition, the *data analysis, statistics, and probability content strand* also had a substantially lower percent of *reproduction* items than the set of NAEP items overall (25 percent compared to 46 percent).

Table 18. Percentage distribution of NAEP 2003 eighth-grade and NAEP 2000 twelfth-grade mathematics problem solving items across PISA competency clusters, by NAEP content strand: 2000 and 2003

PISA competency cluster	Total	NAEP content strand				
		Number sense, properties, and operations	Measurement	Geometry and spatial sense	Data analysis, statistics, and probability	Algebra and functions
Number of NAEP problem solving items[1]	79	17	9	19	16	18
		Percentage distribution				
Reproduction	46	53	56	53	25	44
Connections	44	35	44	37	63	44
Reflection	4	0	0	5	13	0

[1] NAEP data are based on a selected set of 79 mathematics items from the NAEP 2003 grade 8 and NAEP 2000 grade 12 assessments. This selected set of items consists of all extended-response items and all items from the *problem solving* category of the NAEP 2003 framework. All extended-response items are from the *problem solving* category except one item from the *conceptual understanding* category.
NOTE: Percentages reflect the proportion of NAEP items classified by the expert panel to each *competency cluster* in the PISA 2003 mathematical literacy framework. Five NAEP items that the panel did not classify to the PISA framework are not included. Detail may not sum to totals because of rounding or omitted items.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2000 Mathematics Assessment; and Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills*, 2003.

6.3. PISA Situations or Contexts

The *situations or contexts* dimension of the PISA mathematical literacy framework describes the particular context or situation in which a student is engaged in the application of mathematics. These situations are considered to be part of a students' everyday experience within which mathematical tasks are presented and there is a need for mathematical problem solving (OECD 2003). The PISA 2003 framework outlines an ordered taxonomy of four *situations or contexts* based on the "distance" they are from students' experience as explained by the framework. The four *situations or contexts*, starting with the closest to the student, are as follows: *personal, educational/occupational, public,* and *scientific*.

Table 19 compares the distribution of the PISA items and NAEP *problem solving* items across the PISA *situations or contexts* categories. For the set of NAEP items classified by the panel, there is a smaller percentage of items in the *personal* category compared to the PISA items (8 percent compared to 21 percent). There also is a smaller percentage of NAEP items addressing *public situations* (22 percent compared to 34 percent in PISA) and a somewhat greater percentage addressing *educational/occupational situations* (30 percent compared to 24 percent in PISA). A distinction can also be made with regards to the *scientific context*, with 34 percent of the NAEP items compared to 20 percent of PISA items. Consistent with its design, PISA has a reasonably balanced representation across the *situations or contexts*, with a slight emphasis on the *public* category.

Table 19. Percentage distribution of PISA 2003 mathematics items and NAEP 2003 eighth-grade and NAEP 2000 twelfth-grade mathematics problem solving items across PISA situations or contexts categories: 2000 and 2003

PISA situations or contexts category	PISA[1]	NAEP[2]
Personal	21	8
Educational/occupational	24	30
Public	34	22
Scientific	20	34

[1] PISA items classified by PISA developers.
[2] NAEP items classified by expert panel.
NOTE: NAEP data are based on a selected set of 79 mathematics items from the NAEP 2003 grade 8 and NAEP 2000 grade 12 assessments. This selected set of items consists of all extended-response items and all items from the *problem solving* category of the NAEP 2003 framework. All extended-response items are from the *problem solving* category except one item from the *conceptual understanding* category. Percentages reflect the proportion of items classified to each *situations or contexts* category in the PISA 2003 mathematical literacy framework. Five NAEP items that the panel did not classify to the PISA framework, and one PISA item, for which a *situations or contexts* classification was not provided by the assessment developers, are not included. Detail may not sum to totals because of rounding or omitted items.
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2003 Mathematics Assessment; U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2000 Mathematics Assessment; Organization for Economic Cooperation and Development, Program for International Student Assessment (PISA) 2003 Mathematical Literacy Assessment; and Organization for Economic Cooperation and Development (OECD), *Program for International Student Assessment (PISA) 2003 Assessment Framework*, 2003.

6.4. Comparing General Characteristics of NAEP and PISA Items

In addition to classifying NAEP *problem solving* items to the PISA framework dimensions, the panel also evaluated these items with respect to whether or not they could appear on the PISA assessment and, if not, in what way the characteristics of the NAEP items were different from those that might appear on PISA. While some of the NAEP *problem solving* items reviewed by the panel were judged as likely to appear in PISA as they were, a substantial number were identified as not appropriate for PISA. Others were identified as requiring revision in order to be included in the PISA assessment. Some of the general characteristics noted by the panel that distinguished the NAEP and PISA items included the authenticity of the *problem solving context*, the nature of the mathematical application, the level of instructions given, and the general formatting and sequencing of items. The panel noted that scaffolding—breaking an item into component parts and presenting the parts as a series, in steps of increasing complexity to engage students in the task—is used more in NAEP than in PISA.

In general, the NAEP items are presented in isolation, while PISA typically presents a series of items related to a particular *problem solving situation*. In some cases, the panel noted that individual NAEP items might be appropriate for PISA if included as part of a series of items in a larger task. Also, there were other NAEP items that presented a series of questions all within one item, but these were all scored together on a single rubric. In contrast, the panel noted that PISA would present these as separate questions within a larger task, and each would be scored separately.

Some example items are included in appendix E to illustrate some of these general item characteristics noted in comparing NAEP with PISA. Example 23 shows a NAEP *problem solving* item that the panel judged as being appropriate for the PISA assessment. Example 24 illustrates a NAEP *problem solving* item that the panel believed would not appear in the PISA assessment because it is based on a contrived situation. Example 25 was identified as a NAEP problem that might appear on PISA after revision. The panel noted that the presentation of this item would be revised for PISA by breaking it into a series of questions to be assessed and scored separately.

Example 26 shows a PISA graphical interpretation item, which asks students to draw conclusions based on the data in the graph.

This section compared the PISA assessment with NAEP fourth- and twelfth-grade *problem solving* items. The last section includes a summary and conclusion of the findings of this comparison study of NAEP , TIMSS, and PISA.

7. Conclusion

In summary, the content comparisons between NAEP, TIMSS, and PISA reveal some key differences in the mathematics topics covered, grade-level correspondence, and the characteristics of the item pools on other dimensions. All of these factors together may result in differences in student performance, and it is useful to consider these differences when interpreting the results from the three assessments. In addition, the PISA assessment, with its focus on problem solving and the application of mathematics in real-world situations, provides additional information on student performance that is complementary to that obtained from NAEP and TIMSS.

With respect to NAEP and TIMSS, a comparison of the frameworks revealed considerable agreement on the general boundaries and basic organization of mathematics content, with both assessments including five main content areas corresponding to traditional mathematics curricular areas related to *number*, *measurement*, *geometry*, *data*, and *algebra*. Both NAEP and TIMSS frameworks also include dimensions that define a range of cognitive skills and processes that overlap across the two assessments. Despite some apparent similarities at the broadest level, a closer examination of the items in each assessment reveals different emphases placed at the topic and subtopic level, as well as some differences in grade level expectations across mathematics topics. As a result, both NAEP and TIMSS assessments may each contribute more information in some areas as well as some unique components to the larger picture of what students at fourth and eighth grades know and can do in mathematics.

PISA stands apart from NAEP and TIMSS in a number of important areas including a mathematics framework organized around *overarching ideas* rather than curriculum-based content areas, a focus on problem solving in real-world applications, and sampling an age-based population of secondary school students (15-year-olds). Based on the results from this study, PISA also includes larger proportions of constructed-response items and items classified at higher levels of *mathematical complexity* than either NAEP or TIMSS. Although PISA is distinct from NAEP and TIMSS in numerous ways, there are still some similarities when the PISA items are directly compared with the other assessments. The mathematics content covered by most PISA items is consistent with topics included in the NAEP eighth-grade mathematics framework.

This report provides a first-level comparison of items in each assessment in terms of the coverage of broad content areas and distribution across mathematics topics as defined in the frameworks. All items in each assessment were considered in order to make overall comparisons of content coverage and grade-level expectations as well as distributions with respect to three broad levels of mathematical complexity. In addition, the types of item classifications conducted within the time constraints of this study permit comparisons at the mathematics topic level for each content area. While this method provides a broad view of some of the similarities and differences between the assessments, it is limited in terms of the types of comparisons that are provided at the item level. More in depth analyses of the exact nature of the items from each assessment within topics would reveal other important differences related to difficulty, scope, depth, complexity, and other item attributes. These types of more focused comparisons were outside the scope of this study, but may be important to include in future comparative studies of the assessments.

References

Binkley, M., and Kelly, D. (2003). *A Content Comparison of the NAEP and PIRLS Fourth Grade Reading Assessments* (NCES 2003–10). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Binkley, M., and Rust, K. Eds. (1994). *Reading Literacy in the United States: Technical Report of the U.S. Component of the IEA Reading Literacy Study* (NCES 94–259). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Mullis, I.V.S., Martin, M.O., Smith, T.A., Garden, R.A., Gregory, K.D., Gonzalez, E.J., Chrostowski, S.J., and O'Connor, K.M. (2003). *TIMSS Assessment Frameworks and Specifications 2003, 2nd ed*. Chestnut Hill, MA: International Study Center, Lynch School of Education, Boston College.

National Assessment Governing Board (NAGB). (1992). *1996 Mathematics Specifications for the National Assessment of Educational Progress (NAEP)*. Washington, DC: Author.

National Assessment Governing Board (NAGB). (2002). *Mathematics Framework for the 2003 National Assessment of Educational Progress.* Washington, DC: Author.

National Assessment Governing Board (NAGB). (2004). *Mathematics Framework for the 2005 National Assessment of Educational Progress.* Washington, DC: Author.

Neidorf, T.S., Binkley, M., and Stephens, M. (2006). *Comparing Science Content in the National Assessment of Educational Progress (NAEP) 2000 and Trends in International Mathematics and Science Study (TIMSS) 2003 Assessments* (NCES 2006–026). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Nohara, D. (2001). *A comparison of the 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). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

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