

Comparison of the PISA 2022 Mathematics, Reading, and Science Assessments with NAEP

(This page is intentionally left blank)

Comparison of the PISA 2022 Mathematics, Reading, and Science Assessments with NAEP

June 2024

Project Officer
Samantha Burg
National Center for Education Statistics

Maria Stephens
Yemurai Tsokodayi
American Institutes for Research

Contents

Introduction	3
Overview of PISA and NAEP	3
Focus and Scope of the NAEP-PISA Comparison.....	5
Comparison of the PISA and NAEP Mathematics Assessments	7
Mathematics comparison methods	7
How do PISA and NAEP define mathematics?	7
What item response formats and digital functionalities do the PISA and NAEP assessments include?... 9	
Are there mathematics topics that PISA assesses that NAEP does not (and vice versa)?	9
How do PISA and NAEP mathematics items compare?	10
What are some example items that demonstrate these similarities and differences?	11
Comparison of the PISA and NAEP Reading Assessments	16
Reading comparison methods	16
How do PISA and NAEP define reading?	16
What item response formats and digital functionalities do the PISA and NAEP assessments include?. 17	
How do PISA and NAEP reading passages compare?.....	18
How do PISA and NAEP reading items compare?	19
What are some example items that demonstrate these similarities and differences?	20
Comparison of the PISA and NAEP Science Assessments	24
Conclusion.....	25
Appendix A: AIR Experts Supporting the PISA-NAEP 2022 Comparison Study	26
Appendix B: Guide Sheets for Expert Panel Review Meetings.....	27

Introduction

In December 2023, the National Center for Education Statistics (NCES) released [results](#) from the 2022 administration of the Program for International Student Assessment (PISA), an assessment of 15-year-olds’ reading, mathematics, and science literacy. The PISA 2022 release followed the months-earlier release of results from the 2022 National Assessment of Educational Progress (NAEP) [reading](#) and [mathematics](#) assessments at 4th and 8th grades. In both cases, the results were among the first large-scale, multi-subject assessment data on U.S. student performance since prior to the start of the COVID-19 pandemic, heightening public and research interest above their already typically high levels.

This briefing paper provides an overview of these two key NCES data sources: PISA and NAEP. It delves into the similarities and differences between them at a programmatic level and in terms of what content and skills they measure and how they measure them. This information can provide policymakers, researchers, educators, and the public with a deeper understanding of the assessment results, including score trends since the pandemic, and what each contributes to our knowledge of U.S. student performance. The paper is based on a comparison study commissioned by NCES and undertaken by the American Institutes for Research (AIR) in the fall of 2023.

Overview of PISA and NAEP

[PISA](#) is an international assessment that aims to measure students’ cumulative knowledge and skills, acquired both in and outside of school near the end of compulsory education across a range of participating countries. It thus assesses 15-year-old students and focuses on *literacy*—the application of their skills and knowledge—in reading, mathematics, and science. PISA has been conducted every 3 years since 2000, except for a recent one-year delay due to the COVID-19 pandemic which shifted data collection from 2021 to 2022. With each administration of PISA, one of the three core domains is designated as “major,” receiving a longer testing time, and the other two are “minor,” with shorter testing time. Additionally, PISA variously includes optional assessments in subjects such as financial literacy, problem solving, and creative thinking and, since 2015, has transitioned to a digitally-based assessment.

[NAEP](#), which is often called the “Nation’s Report Card,” has measured the academic achievement of a nationally representative sample of students since its inception in 1969. The present study focuses on “main NAEP,” which most regularly assesses 4th-, 8th-, and 12th-grade students in the reading, mathematics, and science they are likely to learn in U.S. classrooms. For 4th- and 8th-grade reading and mathematics, NAEP is generally administered every 2 years (with current trendlines from the 1990s), whereas for 12th grade and for science, the frequency is typically every 4 years (with current trend lines from the 2000s¹). NAEP also periodically assesses other subjects, such as writing, U.S. history, civics, and economics. NAEP not only provides national-level data but, for key subjects, also provides data at the state level and for some [large school districts](#). Finally, alongside main NAEP, [long-term trend NAEP](#) provides data for key age groups in mathematics and reading (not part of the present study). Like PISA, main NAEP has transitioned to digitally-based assessment.

PISA and NAEP thus both serve as important sources for monitoring U.S. student performance over time (exhibit 1), with PISA providing a benchmark for U.S. students’ performance relative to their

¹ This applies to mathematics and science. The 12th-grade reading trendline extends back to 1992.

international peers and NAEP monitoring performance not only at the national level but at subnational levels and among various student groups.

Exhibit 1. PISA and NAEP at a Glance

	PISA	NAEP
Type	International assessment	National assessment
Subjects	Reading literacy, mathematics literacy, and science literacy and <i>optional domains</i> including financial literacy, problem solving, and creative thinking	Reading, mathematics, and science and other subjects including civics, economics, U.S. history, and writing
Target age/grade	15-years-old (U.S. modal grade 10)	4th, 8th, and 12th grades
When current trendlines began	2000 (reading) 2003 (mathematics) 2006 (science)	Early 1990s (reading; 4th- and 8th-grade mathematics) 2005 (12th-grade mathematics) 2009 (science)
Frequency	Every 3 years ¹	Every 2 years ²
Sample size	About 4,600 students sampled	Over 100,000 students sampled per grade/subject
Subnational participation	Some states have participated independently in previous cycles but otherwise no state-level estimates	State-level results for reading, mathematics, and (typically) science; Some district-level results for reading and mathematics ³
Purpose	To measure how well students can apply, by end of compulsory school, their cumulative knowledge and skills to real-life problems and situations	To measure the level of students' academic achievement at the end of key grade spans

¹ There is one exception to this frequency: due to the pandemic, PISA 2021 was delayed to 2022. Additionally, after the 2025 administration, PISA will move to a 4-year cycle.

² This is the frequency for 4th- and 8th-grade reading and mathematics. Every 4 years is the frequency for 12th grade. NAEP was also delayed from 2021 to 2022 due to the pandemic.

³ State- and district-level results are for 4th and 8th grades only. Puerto Rico also regularly participates in NAEP mathematics assessments, and the Department of Defense Education Activity (DoDEA) participates in the mathematics, reading, and (sometimes) science assessments at these grades.

NOTE: Bold indicates the subjects compared in the present study.

As illustrated in exhibit 1, despite assessing the same core subjects, PISA and NAEP differ in terms of their purposes, target populations, and sample sizes. Additional discussion of these and other programmatic differences, including testing windows, measurement precision, and reporting scales, is in exhibit 2.

Exhibit 2. Programmatic Differences Between PISA and NAEP Relevant to the 2022 Results

- **Purpose:** PISA aims to measure students' cumulative knowledge and skills, acquired both in and outside of school, near the end of compulsory education across the range of participating countries. In contrast, NAEP aims to measure academic achievement at the end of specific grade *spans* and is therefore more closely tied to U.S. school curricula.
- **Population(s):** Both PISA and NAEP sample U.S. students in such a way that the assessment results will be generalizable at the national level. However, their specific target populations differ. PISA draws an age-based sample and targets 15-year-old students who, in the United States, are typically in the 10th grade. NAEP draws a grade-based sample and targets students in 4th, 8th, and 12th grades. The differences in approach relate to the differences in each program's purpose—with PISA focusing on the cumulative knowledge and skills of students near the end of compulsory schooling across the participating countries and NAEP being tied more closely to school-based learning in U.S. schools at the end of specific grade spans.
- **Measurement precision:** In 8th-grade mathematics and reading, NAEP reports students' performance at the state level and also for some large school districts. Because of this, NAEP 2022 sampled many more U.S. students than did PISA 2022 (over 100,000 per grade/subject compared to about 4,800 total). Consequently, NAEP measures U.S. students' performance with a higher level of precision (i.e., with considerably smaller standard errors) than PISA and may detect smaller differences than PISA.
- **Testing window:** PISA and NAEP have different testing windows. PISA students are typically in the 10th grade in the United States and tested in the fall semester, whereas the NAEP students closest to the PISA population are in 8th or 12th grade and are tested in the spring semester. PISA 2022 tested students in October and November of 2022, whereas NAEP 2022 tested 8th-grade students in January through March (for mathematics and reading). Twelfth-grade students were last tested in NAEP mathematics and reading in January through March of 2019. This was also the window for the last assessment of NAEP science with 4th-, 8th-, and 12th-grade students.
- **Testing time:** PISA is a slightly longer test than NAEP at around 2 hours of testing time with an additional half-hour for a student questionnaire compared to 1.5 hours, inclusive of the student questionnaire, for NAEP in general. In NAEP, however, students take a single subject whereas in PISA students are tested on multiple subjects (including the major domain and some combination of the other domains).
- **Reporting scales:** PISA's scale ranges from 0 to 1,000 with a standard deviation of about 100 points, whereas NAEP's scales for 8th-grade mathematics and reading (and 12th-grade reading) range from 0 to 500 with standard deviations of about 30 to 40 points. For 12th-grade mathematics the NAEP scale ranges from 0 to 300. As a result, it takes a larger scale-point difference in PISA to approximate a 1-point difference in NAEP.

Focus and Scope of the NAEP-PISA Comparison

Although PISA and NAEP share three subjects (reading, mathematics, and science), the comparison study undertaken by AIR in the fall of 2023 focused primarily on the reading and mathematics assessments. This was because (1) these were the subjects assessed in both PISA and NAEP in 2022, (2) mathematics

was PISA 2022’s major domain and thus had received a framework update, and (3) despite being a minor domain, the PISA 2022 reading assessment was based on a framework that had not been previously studied in comparison to NAEP. For completeness, however, this briefing paper draws from an earlier unpublished comparison study to provide information on how the PISA and NAEP science assessments compare.² In terms of grade focus, the study compared the PISA mathematics and reading assessments with their NAEP counterparts at the most relevant grades: 8th and 12th grades (exhibit 3).

Exhibit 3. Focus of the PISA and NAEP Mathematics and Reading Comparisons

Program	Subject	2019	2022
PISA	Reading	--	15-year-olds
	Mathematics	--	15-year-olds
	Science	--	15-year-olds
NAEP	Reading	Grade 4, Grade 8, Grade 12	Grade 4, Grade 8
	Mathematics	Grade 4, Grade 8, Grade 12	Grade 4, Grade 8
	Science	Grade 4, Grade 8, Grade 12	--

NOTE: The subjects and populations highlighted in bold blue text were incorporated into the fall 2023 comparison study, providing data on both frameworks and items. Information in this paper on grades 8 and 12 science comes from an earlier study comparing PISA 2015 and NAEP 2015 (see footnote 2 below). Grade 4 was not included in these comparisons.

For each subject in the fall 2023 study, a panel of AIR subject matter experts was convened to provide qualitative data on both the frameworks (the documents that guide the development of the assessments) and the items (the stimulus material and test questions given to students).

The **framework review** examined what each assessment intended to measure, with the experts comparing:

- Subject domain definitions and each framework’s descriptions of content and cognitive dimensions; for NAEP this involved review of the framework at both the 8th- and 12th-grade levels
- Targets for percentage distributions of items within the content and cognitive dimensions
- Assessment features, such as item response formats, functionalities included in computer-based tasks, accommodation policies, calculator policy (mathematics), and text authenticity policy (reading)

The **item-to-item comparison** examined how the frameworks were operationalized through the items in order to gain a deeper and more nuanced understanding of what and *how* knowledge and skills are measured and how that might differ between assessments. It involved:

- Clustering seemingly similar items from each assessment for side-by-side examination, both by difficulty and by categories of interest
- Collecting observations about similarities and differences within clusters, using guidance about what to look for as a starting point for discussion

Additional details about methodology are subject specific and described in the sections that follow.

² This study, conducted by AIR in 2016, produced two internal reports related to science: *Science Summary Statements* and *PISA/NAEP/TIMSS Item Comparisons in Science*.

Comparison of the PISA and NAEP Mathematics Assessments

Mathematics comparison methods

The mathematics panel consisted of five content experts with knowledge and experience in mathematics curricula, large-scale assessment generally, and PISA or NAEP specifically (see appendix A). They received training for their role from an AIR senior researcher who served as study coordinator.

For the **framework reviews**, the experts first reviewed the frameworks independently and then participated in a group discussion by webinar to share insights and observations. For PISA, experts reviewed the [PISA 2022 mathematics assessment framework](#), which—because mathematics was the major domain in 2022—was updated from the prior iterations used for the 2003 through 2009 cycles and the 2012 through 2018 cycles. For NAEP, experts reviewed the [NAEP 2022 and 2024 mathematics framework](#), which has been the basis of prior assessments since 2009 (and earlier for 4th and 8th grades).

For the **item-to-item comparison**, a subgroup of three experts first reviewed the full items pools and flagged items according to specific categories of interest (exhibit 4). One expert each was assigned to the PISA 2022, NAEP 2022 grade 8, and NAEP 2019 grade 12 item pools, and they met periodically to ensure they were using a standardized approach.

Exhibit 4. Mathematics Categories of Interest

Items that

1. ask students to create or identify mathematical models
2. ask students to use given mathematical models
3. involve tabular data
4. involve graphical data
5. involve growth phenomena*
6. involve geometric approximation*
7. involve computer simulations*
8. involve conditional decision-making*

* Area of special emphasis in the PISA 2022 mathematics framework/assessment.

The categories of interest included a set of demand- or feature-based categories used in previous comparison studies that proved useful in identifying differences between assessments (1-4 above) and a set of topic-based categories identified as areas of special emphasis in the PISA 2022 framework (5-8 above). These categories were not intended to convey any judgment in what should or should not be included in any assessment, but to provide a variety of lenses with which to observe some more nuanced differences between the two assessments.

Based on the flagging activity, the study coordinator organized clusters according to the 8 categories and, based on student performance data, identified the 5 easiest and 5 most difficult items from each assessment and grade to make a ninth cluster. The expert panel then participated in an in-person meeting to discuss the clusters and provide their insights and observations about similarities and differences between the assessments. They used a guide sheet to help facilitate discussion (see appendix B).

How do PISA and NAEP define mathematics?

Both the PISA and NAEP 2022 frameworks have a content dimension and a cognitive dimension—which together describe what students should know and be able to do and what thinking skills are required. As shown in exhibit 5, **PISA and NAEP differ significantly in their conceptualization of the cognitive**

skills to be measured. They also organize content knowledge somewhat differently, although there is still some overlap within and across each of the content areas. PISA is unique in that *all* items and problem-solving situations are to be in a real-world context whereas NAEP designates a balance between contextualized items and those that are purely mathematical.

Exhibit 5. Overview of the PISA and NAEP Mathematics Frameworks

Dimension	PISA 2022	NAEP 2022
Content knowledge ¹	<p>PISA is organized around <i>four content categories</i> intended to reflect the mathematical phenomena that underlie broad classes of problems, the general structure of mathematics, and the major strands of typical school curricula:</p> <ul style="list-style-type: none"> • Quantity • Space and shape • Change and relationships • Uncertainty and data <p>Within these content categories, the framework also designates <i>four topics for special emphasis</i>:</p> <ul style="list-style-type: none"> • Computer simulations • Geometric approximation • Growth phenomena • Conditional decision-making 	<p>NAEP is organized around <i>five content areas</i>, based on the major disciplinary areas of mathematics:</p> <ul style="list-style-type: none"> • Number properties and operations • Measurement • Geometry • Algebra • Data analysis, statistics, and probability <p>The content areas are further divided into subtopics and objectives to provide a high degree of specificity of the content-related skills targeted by the assessment.</p>
Cognitive skills ²	<p>PISA is organized around <i>mathematical reasoning</i> and <i>three problem-solving processes</i> (formulating situations mathematically; employing mathematical concepts, facts, and procedures; and interpreting, applying, and evaluating mathematical outcomes).</p>	<p>In contrast, NAEP is organized by <i>mathematical complexity</i>, describing three hierarchical levels of cognitive demand (high, moderate, or low), which assumes students’ familiarity with tasks’ mathematics and focuses on what students are asked to do rather than how they are asked to do it.</p>
Other	<p>PISA has <i>context</i> as an explicit organizing dimension (including personal, occupational, societal, and scientific contexts).</p>	<p>NAEP does not have context as an organizing dimension. The framework indicates that the item pool should seek a balance of items that are purely mathematical and items that are set in the contexts of real-world problems.</p>

¹ A [prior study](#) examining the PISA 2012 and NAEP 2013 mathematics frameworks found that, at the content area level, most pairings between PISA and NAEP, respectively (e.g., quantity with number properties and operations; space and shape with measurement; change and relationships with algebra; and uncertainty and data with data analysis, statistics, and probability) were found to be “quite dissimilar with some overlap”. The pairing of PISA space and shape with NAEP geometry was found to be “substantially or wholly different.”

² Per the same study, the cognitive dimensions were found to be “substantially or wholly different” in conceptualization. SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), [PISA 2022 Assessment and Analytical Framework](#); and U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), [Mathematics Assessment Framework for the 2022 and 2024 National Assessment of Educational Progress](#) (also used in 2019).

What item response formats and digital functionalities do the PISA and NAEP assessments include?

Both PISA 2022 and NAEP 2019/2022 include mathematics items that require students to select their responses from a set of choices and those in which they must construct their own responses. Among the latter, both the PISA and NAEP assessments include constructed-response items that require short, discrete answers and those that require longer, more open-ended answers. Based on the target distributions described in their respective frameworks, PISA has a slightly greater emphasis on constructed-response items than NAEP.

Both PISA and NAEP make use of digital functionalities, such as scrolling, navigating between tabs, and the presence of a help menu (exhibit 6). NAEP includes additional unique features, such as the ability to have directions and items read aloud in English through text-to speech, color scheming, and highlighting capabilities. (A list of universal design elements and accommodations that are available in NAEP assessments can be found [here](#) on the NCEs website.) In contrast, PISA has a calculator tool available across items, whereas NAEP has calculators only on identified blocks. Both have some individual tools (e.g., simulators in PISA; rulers and graphing tools in NAEP) that are available for specific items.

Exhibit 6. Example Digital Functionalities at a Glance (Mathematics)

Functionalities	PISA	NAEP
Scrolling	√	√
Navigating	√	√
Help menu	√	√
Text-to-speech		√
Color scheming and highlighting		√
Digital scratch pad		√
Calculator	All items	Only identified blocks
Other tools	Simulators	Rulers, graphing tools, choice eliminator, equation editor

Are there mathematics topics that PISA assesses that NAEP does not (and vice versa)?

In general, PISA 2022 and NAEP 2019/2022 have overlapping mathematics content, with PISA's content categories covering broadly similar ground as NAEP's content areas. However, there are some specific topics that may be more or less commonly reflected in one or the other assessment. Often these relate either to the difference in purpose—PISA's mathematical literacy versus NAEP's curricular achievement—or the fact that NAEP extends to grade 12, covering more advanced topics than PISA, which assesses 15-year-olds typically in 10th grade. For example:

- Topics that were not observed (or were less frequently observed) in PISA items include inequalities on a coordinate plane, graphs of nonlinear functions, normal distributions and boxplots, and exponential equations. PISA also had many fewer algebraic graphs than NAEP. However, because PISA draws from a wide range of real-life sources (such as newspapers), it had a greater range of displays within the graph types in common between the two assessments (except for algebraic). For example, in PISA, similar graph types often appeared visually different depending on the item and context (e.g., stylized versus more traditional), whereas in NAEP,

graphs of the same type were more uniform across the assessment and (by design) in line with the types of graphs encountered in school. In PISA, graphs were also more frequently shown as multiples than in NAEP, with students required to interpret across the graphs. PISA also included circle graphs, which are part of the NAEP framework but were not observed in the given item pools.

- Although NAEP has items that would fit into all four of the “special topics” designated in the PISA 2022 framework (exhibit 5), those related to computer simulation and conditional decision-making may cover a narrower range of specific topics or problem scenarios compared to PISA. This is not necessarily surprising, as these are not explicitly called out in the NAEP framework as they are in the PISA framework. For example, PISA has items where students have to use simulators to test assumptions and explore relationships among variables to solve a problem, whereas such items are not typically found in NAEP. For conditional decision-making, most NAEP items relate to joint probability and are presented in scenarios involving selections of number cards or predicting outcomes of tossing number cubes. Additionally, most NAEP geometry items are focused on finding an answer based on a given figure, or figures, with well-defined values rather than geometric approximations.

How do PISA and NAEP mathematics items compare?

Key similarities in the nature of the items

The PISA 2022 and NAEP 2019/2022 items were similar in some basic features.

- Both PISA 2022 and NAEP 2019/2022 include items set in real-world contexts, including those that are “context dependent” and require students to continually think about how to interpret the context and how to use the context to interpret the results. However, PISA has more of these types of items.
- Both PISA 2022 and NAEP 2019/2022 have items that require students to select a response—either single or multiple responses—and those that require them to construct a response. Both are digitally based assessments and have items enhanced with dynamic capabilities, such as drag-and-drop and spreadsheets that can be populated.
- The easiest items on PISA and NAEP in the most recent assessments were similar in that they tended to lack equations. Additionally, the relatively easy NAEP items often included a visual aspect and/or were single step, and in PISA, they appeared to be on the lower end of the reading load. The most difficult items on PISA were multi-step. On NAEP, they required reasoning, presented students with equations, were not set in a real-world context, or required abstract thinking.

Key differences in the nature of the items

While there is a general overlap in content and some features, PISA 2022 and NAEP 2019/2022 items can appear quite dissimilar in certain cases. Some of these dissimilarities are related to:

- *Context.* As mentioned above, NAEP includes items set in real-world problem-solving contexts, including those where the context plays a significant role in the student’s thinking. It also includes—per its framework—purely mathematical items. In contrast, all of PISA’s items are contextualized. Among these contextualized items, PISA tends to have more text supporting and

describing the context; the contexts may be more authentic and wide-ranging; and there may be more associated visual material. In contrast, NAEP contexts are more streamlined, use more simplified language, and are more closely tied to the types of problem-solving scenarios that students are likely to encounter in school. NAEP also avoids contexts that overlap too strongly with other content areas, such as science, whereas PISA intentionally includes such contexts.

- *Reading load.* As noted in the previous bullet, NAEP uses as simple language as possible to describe problem scenarios so that students' reading skills do not influence their mathematics performance. In contrast, PISA aims for highly contextualized scenarios, and its units thus often require students to read and sift through more information. Additionally, all NAEP items have a read-aloud feature to reduce the reading load.
- *Directedness and discreteness.* Because the NAEP framework specifies grade-level assessment objectives, many items focus on discrete knowledge or skills. Thus, NAEP items are more commonly prescriptive of a problem-solving strategy, requiring students to respond based on a particular system of thinking. In contrast, some PISA items demand more exploration on the part of students in both setting up and following through on a problem-solving process. For example, while NAEP has items that require students to either create or identify a mathematical model or use a model, it is less common to find both aspects combined in a single item. PISA, however, does commonly have such items.
- *Item grouping.* PISA mathematics items are arranged in units that share stimulus material. Within a unit, there are usually multiple items, involving different mathematical knowledge or skills, that relate to the common scenario. While there are some mathematics items grouped in this way in NAEP, these are an exception rather than the general rule.

What are some example items that demonstrate these similarities and differences?

The following pages present four pairs of example items that demonstrate some of the similarities and differences just discussed. These examples are drawn from the released item sets from PISA and NAEP 2022, as well as prior years, and they are from seemingly similar content areas and subtopics. Note that the PISA items tend to be single items from within a larger multi-item unit, whereas the NAEP items are stand-alone. Descriptive information on the PISA items may reference proficiency levels, which in mathematics range from the lowest (1c) to the highest (6), with students reaching a particular level by correctly answering a majority of items at that level.

Item Pair 1

This illustration contrasts a PISA Uncertainty and Data item with a NAEP Data Analysis, Statistics, and Probability item (grade 12). Both items require students to understand graphical data, although because the PISA item is complex multiple choice, it requires three interpretations for full credit. (Other NAEP items may have a similar item response format.) As is typical, the PISA item has a heavier reading load and a potentially less familiar context than does the NAEP item. In contrast, the use of a scatterplot, as in the NAEP item, would be less common in PISA.

PISA

Unit and number: DVD Sales (CMA106 Q1 of 3)
 Item type: Complex multiple choice
 Cognitive classification: Integrate/evaluate
 Difficulty: Proficiency level 4 (full credit); 1a (partial credit)

PISA 2022

DVD Sales
Question 1 / 3

Refer to "DVD Sales" on the right. Click on the choices in the table to answer the question.

Are the statements in the table below supported by the information shown in the graph? Click on either **Yes** or **No** for each statement.

Statement	Yes	No
The number of DVDs sold declined by about 50% from 2008 through 2014.	<input type="radio"/>	<input type="radio"/>
The number of DVDs sold declined by the same amount each year from 2008 through 2014.	<input type="radio"/>	<input type="radio"/>
The slope of the line is the average yearly decline in DVDs sold from 2008 to 2014.	<input type="radio"/>	<input type="radio"/>

DVD SALES

The total number of DVDs sold each year in the United Kingdom from 2008 to 2014 is shown in the graph below.

The values on the horizontal axis represent the number of years **after** 2008. Place the cursor over the points in the graph to see the coordinates of that point. For example, the point (0, 252.9) indicates that there were 252.9 million DVDs sold in 2008. The point (1, 234.6) indicates that there were 234.6 million DVDs sold during the year 2009, etc.

A line has been added to the graph to model these data points.

SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2022 Mathematics Released Items (Field Trial).
<https://pisa2022-questions.oecd.org/platform/index.html?user=&unit=MAT/MA106-DVDSales&lang=eng-ZZZ>

NAEP (Grade 12)

Unit and number: M189601
 Item type: Simple multiple choice
 Cognitive classification: Low complexity
 Difficulty: Easy

In the scatterplot below, each point represents the adult heights of a sister-brother pair.

SISTER-BROTHER HEIGHTS

What is the brother's height that corresponds to the tallest sister's height?

- A. 74 inches
- B. 71 inches
- C. 67 inches
- D. 66 inches
- E. 62 inches

SOURCE: National Center for Education Statistics (NCES), National Assessment of Educational Progress (NAEP), *NAEP Questions Tool*, 2013 Grade 12 Mathematics Assessment.
<https://www.nationsreportcard.gov/nqt/searchquestions>

Item Pair 2


This illustration contrasts a PISA Change and Relationships item with a NAEP Algebra item (grade 8). Both items ask students to identify a model in simple multiple-choice format but, again, the PISA item provides more information to sift through and an extended scientific scenario that would not be prevalent in NAEP.

PISA
Unit and number: Penguins (PM921 Q3 of 4) Item type: Simple multiple choice Cognitive classification: Formulate Difficulty: Not given
Jean assumes the colony will continue to grow in the following manner:
<ul style="list-style-type: none">At the beginning of each year, the colony consists of equal numbers of male and female penguins who form couples.Each penguin couple raises one chick in the spring of each year.By the end of each year 20% of all the penguins (adults and chicks) will die.One year old penguins will also raise chicks.
Based on the above assumptions, which of the following formulae describes the total number of penguins, P , after 7 years?
A $P = 10\,000 \times (1.5 \times 0.2)^7$ B $P = 10\,000 \times (1.5 \times 0.8)^7$ C $P = 10\,000 \times (1.2 \times 0.2)^7$ D $P = 10\,000 \times (1.2 \times 0.8)^7$
<i>NB: Student knows the starting colony size of 10,000 from the prior item.</i>
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2012 Released Mathematics Items. https://www.oecd.org/pisa/pisaproducts/pisa2012-2006-rel-items-maths-ENG.pdf (p. 56)

NAEP (Grade 8)
Unit and number: M414001 Item type: Simple multiple choice Cognitive classification: Low complexity Difficulty: Hard
Sara plans to exercise for 315 minutes over a period of 7 days. She is going to exercise an equal number of minutes each day.
Which of the following equations represents the number of minutes, m , Sara has exercised after d days?
A <input type="radio"/> $m = 45d$
B <input type="radio"/> $m = 315d$
C <input type="radio"/> $m = 315 + d$
D <input type="radio"/> $m = \frac{315}{d}$
E <input type="radio"/> $m = \frac{45}{d}$
SOURCE: National Center for Education Statistics (NCES), National Assessment of Educational Progress (NAEP), <i>NAEP Questions Tool</i> , 2022 Grade 8 Mathematics Assessment. https://www.nationsreportcard.gov/nqt/searchquestions

Item Pair 3

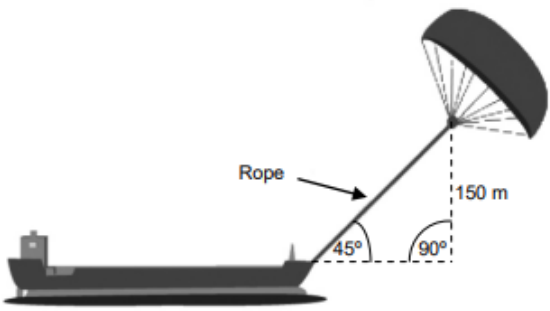
This illustration contrasts a PISA Change and Relationships item with a NAEP Algebra item (grade 12). The PISA item employs a constructed-response format, which is more common in PISA, and includes an illustration that provides no additional information related to the mathematics needed to solve the problem, which would be less likely in NAEP. The NAEP item, in contrast, is purely mathematical, which would not occur in PISA.

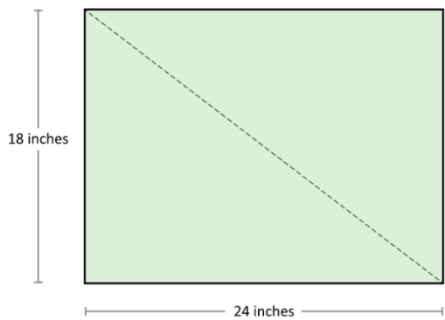
PISA	
Unit and number: Drip Rate (PM903 Q1 of 3) Item type: Extended constructed response Cognitive classification: Employ Difficulty: Not given	
Infusions (or intravenous drips) are used to deliver fluids and drugs to patients.	
	<p>A nurse wants to double the time an infusion runs for.</p> <p>Describe precisely how D changes if n is doubled but d and v do not change.</p> <p>.....</p> <p>.....</p> <p>.....</p>
<p>Nurses need to calculate the drip rate, D, in drops per minute for infusions.</p> <p>They use the formula $D = \frac{dv}{60n}$ where</p> <p>d is the drop factor measured in drops per millilitre (mL)</p> <p>v is the volume in mL of the infusion</p> <p>n is the number of hours the infusion is required to run.</p>	
<p><i>NB. The left-hand image is the item introduction. The right-hand image is the first item in the unit.</i></p>	
<p>SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2012 Released Mathematics Items. https://www.oecd.org/pisa/pisaproducts/pisa2012-2006-rel-items-maths-ENG.pdf (pp. 6–7)</p>	

NAEP (Grade 12)	
Unit and number: M181601 Item type: Simple multiple choice Cognitive classification: Low complexity Difficulty: Hard	
<p>Which of the following expressions is equal to $\frac{1}{x+2} - \frac{2}{x+1}$?</p>	
<p>A. $\frac{-1}{2x+3}$</p> <p>B. $\frac{-x-3}{x^2+2}$</p> <p>C. $\frac{-1}{x^2+3x+2}$</p> <p>D. $\frac{-x-3}{x^2+3x+2}$</p> <p>E. $\frac{-x+5}{x^2+3x+2}$</p>	
<p>SOURCE: National Center for Education Statistics (NCES), National Assessment of Educational Progress (NAEP), 2009 <i>NAEP Questions Tool</i>, 2009 Grade 12 Mathematics Assessment. https://www.nationsreportcard.gov/nqt/searchquestions</p>	

Item Pair 4

This illustration contrasts a PISA Space and Shape item with a NAEP Geometry item (grade 8). Both items require students to use the Pythagorean Theorem to find or estimate the value of a hypotenuse, and both are contextualized, though PISA is again in a scientific context and NAEP is in one typical for school mathematics. PISA uses a nonscaled illustration, whereas NAEP uses a scaled diagram. PISA requires students to recognize that the triangle is isosceles and allows them to approximate the answer.

PISA	
Unit and number: Sailing ships (PM923 Q2 of 4) Item type: Simple multiple choice Cognitive classification: Employ Difficulty: Not given	
<p>Question 3: SAILING SHIPS</p> <p>Approximately what is the length of the rope for the kite sail, in order to pull the ship at an angle of 45° and be at a vertical height of 150 m, as shown in the diagram opposite?</p> <p>A. 173 m B. 212 m C. 285 m D. 300 m</p>	 <p>Note: Drawing not to scale. © by skysails</p>
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2012 Released Mathematics Items. https://www.oecd.org/pisa/pisaproducts/pisa2012-2006-rel-items-maths-ENG.pdf (p. 13)	

NAEP (Grade 8)	
Unit and number: M388201 Item type: Short constructed response Cognitive classification: Low complexity Difficulty: Hard	
	 <p>Mary has a rectangular poster board that measures 18 inches by 24 inches.</p> <p>What is the length of the diagonal of the poster board, in inches?</p> <p><input type="text"/> inches</p>
SOURCE: National Center for Education Statistics (NCES), National Assessment of Educational Progress (NAEP), <i>NAEP Questions Tool</i> , 2022 Grade 8 Mathematics Assessment https://www.nationsreportcard.gov/nqt/searchquestions	

Comparison of the PISA and NAEP Reading Assessments

Reading comparison methods

Like the mathematics comparison, the reading comparison consisted of a framework review and item-to-item comparisons. The reading study also specifically compared the passages that students are asked to read in the assessments. For these activities, AIR convened an in-person expert panel, which consisted of two reading content experts with knowledge and experience with the reading curricula and large-scale assessments, including NAEP (see appendix A). They received training for their role from an AIR senior researcher who served as study coordinator and who facilitated the meeting.

For the **framework review**, the experts were presented with an overview of the two assessments, focusing on their purpose and features, definitions of reading, and organizing dimensions, including content and cognitive processes. The panel then read over the content dimensions of both the PISA and NAEP assessments, highlighting key similarities in how the features of passages/texts were described within each category. For PISA, the focus was the [PISA 2018 reading assessment framework](#), which was updated at that time from prior iterations and was used again in 2022. For NAEP, the focus was the [NAEP 2022 and 2024 reading framework](#), which has been the basis of assessments since 2009.

For the **item-to-item and passage comparisons**, the study coordinator identified analogous content categories based on the framework review and discussion (e.g., narration in PISA and literary fiction/nonfiction in NAEP, exposition in both). The panel then reviewed a sample of PISA and NAEP passages within the analogous categories and within those that were more distinct, along with their related items, to identify more nuanced similarities and differences. The panel used a guide sheet of potential variations to consider during their discussion (see appendix B). The panel also reviewed clusters of items based on their response format, cognitive requirements, and other features, as well as clusters of the three easiest and three hardest items from each assessment based on student performance.

How do PISA and NAEP define reading?

Both the PISA and NAEP 2022 reading frameworks provide a definition of reading, as well as describe what students should read (content dimension) and what thinking skills are required (exhibit 7). **The PISA and NAEP reading frameworks are more similar to each other than are the PISA and NAEP mathematics frameworks, especially concerning the cognitive processes to be assessed and several overlapping passage types. NAEP is unique in its additional focus on assessing “meaning vocabulary,” its inclusion of poetry texts, and its use of authentic, unaltered texts. PISA, in contrast, may have shorter, more transactional, and noncontinuous material, as well as material that has been altered for the purposes of the assessment.**

Exhibit 7. Overview of the PISA and NAEP 2022 Reading Frameworks

	PISA 2022	NAEP 2022
Definition of reading	Reading literacy is understanding, using, evaluating, reflecting on, and engaging with texts in order to achieve one’s goals to develop one’s knowledge and potential and to participate in society.	Reading is an active and complex process that involves: <ul style="list-style-type: none"> • understanding written text, • developing and interpreting meaning, and • using meaning as appropriate to type of text, purpose, and situation.
Content knowledge (<i>what students read</i>)	PISA is organized around: <ul style="list-style-type: none"> • Text type (argumentation, exposition, description, narration, instruction, transactional) • Text format (continuous, noncontinuous, mixed) • Source (single multiple) • Organizational and navigational structure (static, dynamic) 	NAEP is organized around two types of texts : <ul style="list-style-type: none"> • Literary (fiction, literary nonfiction, and poetry) • Informational (exposition, argumentation and persuasion, and procedural texts and documents)
Cognitive processes	PISA is organized around three text processing processes (locating information, understanding, and evaluating and reflecting).	NAEP is organized around three cognitive targets (locate and recall, integrate and interpret, and critique and evaluate).
Other	PISA has situation as an explicit organizing dimension (including personal, educational, occupational, public).	NAEP does not have situation, or context, as an organizing dimension. However, the framework does indicate that passages should represent practical, academic, and other contexts from grade-appropriate sources spanning content areas. NAEP also assesses meaning vocabulary , or how well students apply their understanding of word meanings to passage comprehension.

SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), [PISA 2018 Assessment and Analytical Framework](#) (also used in 2022); and U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), [Reading Assessment Framework for the 2022 and 2024 National Assessment of Educational Progress](#) (also used in 2019).

What item response formats and digital functionalities do the PISA and NAEP assessments include?

Both PISA 2022 and NAEP 2019/2022 include reading items that require students to select their responses from a set of choices and those in which they must construct their own responses. Among the latter, both the PISA and NAEP assessments include constructed-response items that require short, discrete answers and those that require longer, more open-ended answers.

Both PISA and NAEP make use of digital functionalities, such as scrolling, navigating between tabs, and the presence of a help menu (exhibit 8). NAEP includes additional unique features, such as the ability to have directions read aloud in English through text-to-speech, color scheming, and highlighting

capabilities. A list of universal design elements and accommodations that are available on NAEP assessments can be found [here](#) on the NCES website.

Exhibit 8. Example Digital Functionalities at a Glance (Reading)

Functionalities	PISA	NAEP
Scrolling	√	√
Navigating	√	√
Help menu	√	√
Text-to-speech instructions		√
Color scheming and highlighting		√
Lookback buttons		√

How do PISA and NAEP reading passages compare?

Key similarities between PISA and NAEP passages

Both PISA 2022 and NAEP 2019/2022 include a wide range of text types in their assessment, with significant overlap.

- PISA’s narration text type—which includes excerpts from novels, short stories, plays, biographies, comic strips, and newspaper reports of events—overlaps with the fiction and literary nonfiction text types in NAEP.
- Additionally, PISA’s exposition, argumentation, and instruction text types have significant overlap with NAEP’s exposition, argumentation and persuasion, and procedural texts and documents text types, respectively.

Additionally, although the NAEP 2022 reading framework does not explicitly classify texts by as many characteristics as does PISA (e.g., structure, format), both assessments still make use of different organization and navigational structures (e.g., multiple screen pages), as well as different text formats. Both also make use of both single- and multiple-source texts, although PISA does so to a greater extent.

Key differences between PISA and NAEP passages

Despite significant overlap in the type of texts included in the PISA 2022 and NAEP 2022 reading assessments, there are some important differences.

- A major differentiator is that PISA texts specifically include examples of web-based or digital text formats, focusing on texts that students may encounter in their daily lives both in and outside of school (e.g., chats, text messages, online reviews, websites). In contrast, NAEP passages are more reflective of texts that students would encounter in a classroom setting.
- Another key differentiator is that the NAEP 2022 framework emphasizes the authenticity of texts and notes a commitment to selecting high-quality, authentic stimulus materials that students are likely to encounter both in school and outside of school. This results in NAEP having more continuous passages with very few edits to the original text, as well as passages with minimum length requirements (i.e., the grade 8 word count range is 400–1,000 words; the grade 12 range is

500–1,500). In contrast, PISA does not have such requirements and generally has more noncontinuous texts with lower word counts.³

- Finally, NAEP includes poetry passages, whereas PISA does not.

How do PISA and NAEP reading items compare?

Key similarities in the nature of the items

- PISA and NAEP generally have similar item types (multiple choice and constructed response).
- Both PISA and NAEP include item sets—called units in PISA and blocks in NAEP—that are based on single texts and those that are based on multiple texts (up to two in NAEP, two or more in PISA).
- The easiest items on PISA and NAEP have similar characteristics. On both assessments, easy items tend to involve students locating and recalling information almost verbatim from the text. Some NAEP meaning vocabulary items were also among the easy items.
- The most difficult items in PISA and NAEP have similar features in that they are generally constructed-response items that ask students to (1) look across multiple texts for answers, (2) select a perspective/opinion, and (3) provide evidence from the texts to support their perspective/opinion.

Key differences in the nature of the items

- PISA includes items that sometimes allow students to answer questions using knowledge gained outside school without requiring specific reference to or inference based on the passage. In contrast, NAEP items only require responses based on the passages. Because they are curriculum-based, NAEP items are also unique in that students must have knowledge of specific literary devices (e.g., poetry, themes, irony) to answer questions. PISA generally includes fewer items per unit (3–7 items) compared to NAEP blocks (8–11 items), regardless of whether the unit/block includes a single text or multiple texts.
- NAEP includes two types of constructed-response items (short and extended), while PISA constructed-response items generally require shorter responses.
- NAEP also has unique item formats where students can choose their answers directly from highlighted text in the passage.

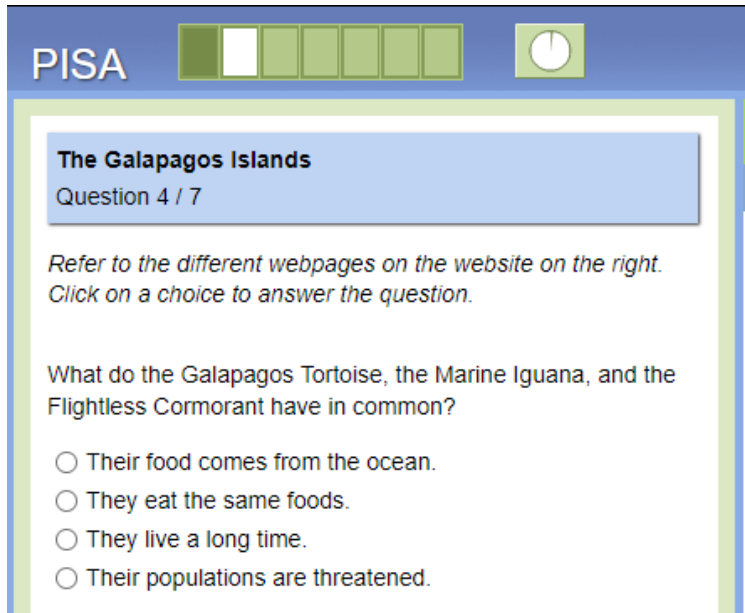
³ The present study did not include passage analyses such as word counts or readability analyses. The last study of [PISA and NAEP reading in 2009](#) did confirm a lower average word count for PISA passages (354, with a wide range) versus NAEP passages (974 for grade 8 and 1,174 for grade 12). A more recent (unpublished) analysis of PISA 2015 reading passages confirmed a similar average word count in the mid-300s. The most recent framework changes in the PISA reading framework are not likely to have induced changes impacting this general comparison.

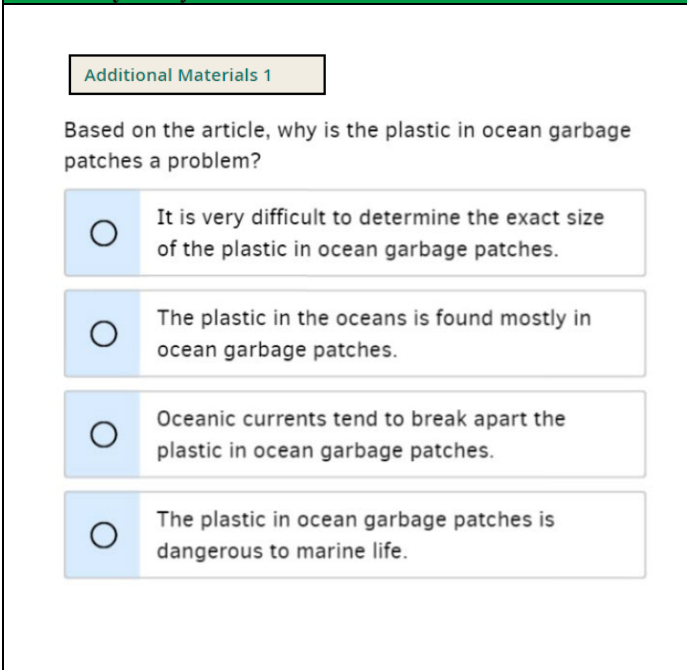
What are some example items that demonstrate these similarities and differences?

The following pages present three pairs of example items that demonstrate some of the similarities and differences just discussed. These examples are drawn from the released item sets from PISA and NAEP 2022, as well as prior years. Note that the descriptive information on the PISA items may reference proficiency levels, which in reading range from the lowest (1c) to the highest (6), with students reaching a particular level by correctly answering a majority of items at that level.

Item Pair 1

Item Pair 1 illustrates an example of **similar** reading item response formats in PISA and NAEP. Both items are simple multiple-choice items where students select one of four response options to answer the question. In this case, both items require students to integrate information from the text as a whole to select the correct response.

PISA	
Unit: Galapagos Islands Item Number: Question 4 of 7 (CR571Q06) Item Type: Simple Multiple Choice Classification: Integrate and Generate Inferences (Understand) Difficulty: Not Provided	
 <p>The screenshot shows the PISA assessment interface. At the top, it says 'PISA' and has a progress bar with 7 segments, the first of which is filled. Below that, the title 'The Galapagos Islands' and 'Question 4 / 7' are displayed. The instruction reads: 'Refer to the different webpages on the website on the right. Click on a choice to answer the question.' The question is: 'What do the Galapagos Tortoise, the Marine Iguana, and the Flightless Cormorant have in common?' There are four radio button options: 'Their food comes from the ocean.', 'They eat the same foods.', 'They live a long time.', and 'Their populations are threatened.'</p>	
SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2018 Reading Assessment. https://pisa2018-questions.oecd.org/platform/index.html?user=&domain=REA&unit=R571-Galapagos&lang=eng-ZZZ	

NAEP (Grade 8)	
Block: Oceans of Plastic Item Number: Question 1 of 8 (Question ID:2022-8R71 #1 R092301) Item Type: Multiple Choice Classification: Integrate/Interpret Difficulty: Easy	
 <p>The screenshot shows the NAEP assessment interface. At the top, it says 'Additional Materials 1'. Below that, the question is: 'Based on the article, why is the plastic in ocean garbage patches a problem?' There are four radio button options: 'It is very difficult to determine the exact size of the plastic in ocean garbage patches.', 'The plastic in the oceans is found mostly in ocean garbage patches.', 'Oceanic currents tend to break apart the plastic in ocean garbage patches.', and 'The plastic in ocean garbage patches is dangerous to marine life.'</p>	
SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), <i>NAEP Questions Tool</i> , 2022 Grade 8 Reading Assessment. https://www.nationsreportcard.gov/nqt/searchquestions	

Item Pair 2

Item Pair 2 displays an example of **similar** reading item response formats in PISA and NAEP. Both items are constructed-response items where students select a perspective from the response options, then provide justification for their selection using information from the text(s).

PISA

Unit: Cow's Milk
Item Number: Question 7 of 7 (CR571Q06)
Item Type: Open response – Human coded
Classification: Detect and handle conflict (Evaluate and reflect)
Difficulty: 506 – Level 3

PISA 🔍

Cow's Milk
Question 7 / 7

Refer to both sources on the right by clicking on each of the tabs. Click on a choice and then type an explanation to answer the question.

Anna, Christopher and Sam are talking about the two texts.

Christopher: No matter what the coffee shop owner does, I'm going to keep drinking milk every day. It's really good for you.

Anna: Not me! I'm going to drink a lot less milk from now on if it's not good for you.

Sam: I don't know, I think we need to know more before we make a conclusion.

With whom do you agree?

Christopher
 Anna
 Sam

Explain your answer. Refer to information from at least one of the texts.

SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2018 Reading Assessment. <https://pisa2018-questions.oecd.org/platform/index.html?user=&domain=REA&unit=R557-CowsMilk&lang=eng-ZZZ>

NAEP (Grade 8)

Block: Oceans of Plastic
Item Number: Question 4 of 8 (Question ID:2022-8R71 #4 R092304)
Item Type: Short Constructed Response
Classification: Critique/Evaluate
Difficulty: Hard

Additional Materials 1Additional Materials 2

Help Diana and Alex decide whether to use one or both photos.

Choose the decision that you think will best help site visitors understand what plastic in ocean garbage patches looks like.

Use photo A only

Use photo B only

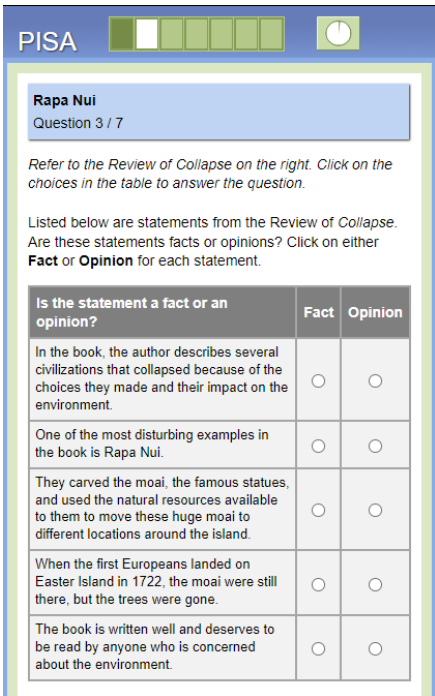
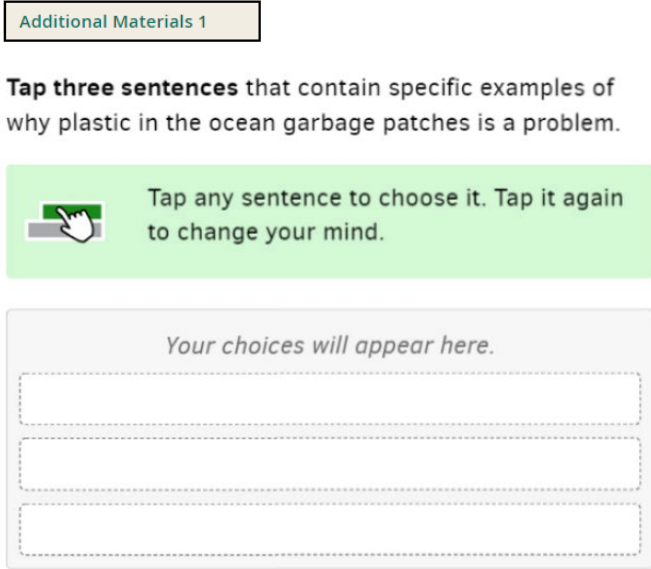
Use both photo A and photo B

Explain your choice using information from the article.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), *NAEP Questions Tool*, 2022 Grade 8 Reading Assessment. <https://www.nationsreportcard.gov/nqt/searchquestions>

Item Pair 3

Item Pair 3 displays an example of reading item response formats that were **observed only in either PISA or NAEP**. The PISA assessment includes items where students answer a series of dichotomous questions (e.g., “yes/no,” “true/false,” “fact/opinion”). The NAEP assessment includes items where students respond to an item by tapping sentences in the text to provide their response.

PISA	NAEP (Grade 8)																		
<p>Unit: Rapa Nui Item Number: Question 3 of 7 (CR551Q06) Item Type: Complex Multiple Choice – Computer Scored Classification: Reflect on content and form (Evaluate and Reflect) Difficulty: 654 – Level 5</p>	<p>Block: Oceans of Plastic Item Number: Question 7 of 8 (Question ID:2022-8R71 #7 R0923MS) Item Type: Selected Response Classification: Locate/Recall Difficulty: Hard</p>																		
 <p>The screenshot shows a PISA assessment interface. At the top, it says 'PISA' with a progress bar and a moon icon. Below that, it says 'Rapa Nui Question 3 / 7'. The instructions are: 'Refer to the Review of Collapse on the right. Click on the choices in the table to answer the question.' Below the instructions, it says: 'Listed below are statements from the Review of Collapse. Are these statements facts or opinions? Click on either Fact or Opinion for each statement.'</p> <table border="1" data-bbox="451 862 831 1252"> <thead> <tr> <th>Is the statement a fact or an opinion?</th> <th>Fact</th> <th>Opinion</th> </tr> </thead> <tbody> <tr> <td>In the book, the author describes several civilizations that collapsed because of the choices they made and their impact on the environment.</td> <td style="text-align: center;"><input type="radio"/></td> <td style="text-align: center;"><input type="radio"/></td> </tr> <tr> <td>One of the most disturbing examples in the book is Rapa Nui.</td> <td style="text-align: center;"><input type="radio"/></td> <td style="text-align: center;"><input type="radio"/></td> </tr> <tr> <td>They carved the moai, the famous statues, and used the natural resources available to them to move these huge moai to different locations around the island.</td> <td style="text-align: center;"><input type="radio"/></td> <td style="text-align: center;"><input type="radio"/></td> </tr> <tr> <td>When the first Europeans landed on Easter Island in 1722, the moai were still there, but the trees were gone.</td> <td style="text-align: center;"><input type="radio"/></td> <td style="text-align: center;"><input type="radio"/></td> </tr> <tr> <td>The book is written well and deserves to be read by anyone who is concerned about the environment.</td> <td style="text-align: center;"><input type="radio"/></td> <td style="text-align: center;"><input type="radio"/></td> </tr> </tbody> </table>	Is the statement a fact or an opinion?	Fact	Opinion	In the book, the author describes several civilizations that collapsed because of the choices they made and their impact on the environment.	<input type="radio"/>	<input type="radio"/>	One of the most disturbing examples in the book is Rapa Nui.	<input type="radio"/>	<input type="radio"/>	They carved the moai, the famous statues, and used the natural resources available to them to move these huge moai to different locations around the island.	<input type="radio"/>	<input type="radio"/>	When the first Europeans landed on Easter Island in 1722, the moai were still there, but the trees were gone.	<input type="radio"/>	<input type="radio"/>	The book is written well and deserves to be read by anyone who is concerned about the environment.	<input type="radio"/>	<input type="radio"/>	 <p>The screenshot shows a NAEP assessment interface. At the top, it says 'Additional Materials 1'. Below that, it says: 'Tap three sentences that contain specific examples of why plastic in the ocean garbage patches is a problem.' Below the instructions, there is a green box with a hand icon and the text: 'Tap any sentence to choose it. Tap it again to change your mind.' Below that, there is a grey box with the text: 'Your choices will appear here.' and three dashed boxes for responses.</p>
Is the statement a fact or an opinion?	Fact	Opinion																	
In the book, the author describes several civilizations that collapsed because of the choices they made and their impact on the environment.	<input type="radio"/>	<input type="radio"/>																	
One of the most disturbing examples in the book is Rapa Nui.	<input type="radio"/>	<input type="radio"/>																	
They carved the moai, the famous statues, and used the natural resources available to them to move these huge moai to different locations around the island.	<input type="radio"/>	<input type="radio"/>																	
When the first Europeans landed on Easter Island in 1722, the moai were still there, but the trees were gone.	<input type="radio"/>	<input type="radio"/>																	
The book is written well and deserves to be read by anyone who is concerned about the environment.	<input type="radio"/>	<input type="radio"/>																	
<p>SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), 2018 Reading Assessment. https://pisa2018-questions.oecd.org/platform/index.html?user=&domain=REA&unit=R551-RapaNui&lang=eng-ZZZ</p>	<p>SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), <i>NAEP Questions Tool</i>, 2022 Grade 8 Reading Assessment. https://www.nationsreportcard.gov/nqt/searchquestions</p>																		

Comparison of the PISA and NAEP Science Assessments

Although PISA 2022 included a science literacy assessment, NAEP last assessed science in 2019. Despite the lack of data from comparable time periods, it is informative to examine how these two assessments compare. Because the PISA 2022 science assessment was based on a framework last updated in [2015](#), and the NAEP 2019 science assessment was based on a framework in use since 2009, we can look to an earlier unpublished study that compared the 2015 PISA and 5 NAEP cycles for insights (see earlier footnote 2).

Based on that study, the PISA and NAEP science frameworks were found to differ in their degree of specificity and directedness. Within its three broad content areas, NAEP defines the topics, subtopics, and grade-specific assessment objectives to be covered. In contrast, PISA provides a broad set of exemplar topics in each content knowledge area, and items developed for PISA may require knowledge in one of the broad topics but are not restricted to that content. In both cases, the desired content can be covered in any combination with any of the cognitive categories (competencies in PISA and practices in NAEP).

Notably, all of PISA’s exemplar topics are covered in the NAEP framework. However, PISA lacks a competency directly corresponding to the NAEP science practice of Identifying Science Principles. In NAEP, this is a practice that draws on declarative knowledge and is used as part of the foundational knowledge necessary for engaging with inquiry and it slowly diminishes in the assessment over the grades, being replaced by an increased focus on Using Scientific Principles and Inquiry. Still, PISA appears to place more emphasis on competencies related to scientific inquiry (exhibit 9); NAEP does emphasize scientific inquiry but increasingly at the upper grades while also providing opportunities for lower performing students.

Exhibit 9. Overview of PISA and NAEP Science Frameworks

Dimension	PISA 2022	NAEP 2019
Content knowledge	PISA is organized around <i>three types of knowledge</i> : <ul style="list-style-type: none"> • Content • Procedural • Epistemic Within <i>content knowledge</i> , PISA focuses on: <ul style="list-style-type: none"> • Physical systems • Living systems • Earth and space systems 	NAEP is organized around <i>three content areas</i> : <ul style="list-style-type: none"> • Physical Science • Life Science • Earth and Space Science
Cognitive skills	PISA is organized around <i>three scientific competencies</i> : <ul style="list-style-type: none"> • Explain phenomena scientifically • Evaluate and design scientific inquiry • Interpret data and evidence scientifically 	NAEP is organized around <i>four science practices</i> : <ul style="list-style-type: none"> • Identifying Science Principles • Using Science Principles • Using Scientific Inquiry • Using Technological Design

Other	PISA has <i>context</i> as an explicit organizing dimension in which to situate the content knowledge and cognitive skills it aims to measure (including personal, local/national, and global).	NAEP does not have situation, or context, as an organizing dimension although individual items are developed that involve local, regional, or global issues.
-------	---	--

NOTE: SOURCE: Organization for Economic Cooperation and Development (OECD), Program for International Student Assessment (PISA), [PISA 2015 Assessment and Analytical Framework](#); and U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), [Science Assessment Framework for the 2019 National Assessment of Educational Progress](#).

Science will be the major domain in PISA’s next cycle in 2025 and thus have an [updated framework](#). For NAEP, [a new framework](#) was adopted in November 2023 to serve as the basis of the 2028 science assessment.

Conclusion

The PISA and NAEP assessments in mathematics, reading, and science are constructed based on their individual respective assessment frameworks. While there is a **general overlap in knowledge and skills between PISA and NAEP** for each subject, the differing frameworks lead to **differences in content emphases, the inclusion of content unique to one or the other, and some items that appear fairly dissimilar** between the assessments.

- Among the subjects PISA and NAEP have in common, the reading assessments have the greatest overlap, especially in terms of the cognitive skills being assessed. The differences in the reading assessments are mainly in the types of passages that students read—with PISA including the type of web-based or digital texts that students may encounter in their daily lives and NAEP having more continuous passages, more school-based texts, and poetry.
- The PISA frameworks in mathematics and science specify broad content areas similar to those in NAEP, though they are less precisely defined in terms of specific assessment content objectives.
- In mathematics and science, there are differences in how the required cognitive skills are defined and, in mathematics particularly, differences in the nature of the items owing to PISA’s strong focus on contextualization.

Ultimately, many of the differences between PISA and NAEP are rooted in the literacy focus of the former compared to the curricular focus of the latter. This difference in purpose influences the types of passages used in reading and the nature and degree of contextualization in mathematics and science problem solving scenarios and items, which distinguishes the PISA and NAEP items and thus the overall assessments from one another.

Appendix A: AIR Experts Supporting the PISA-NAEP 2022 Comparison Study

Mathematics

Audrey Altieri

Researcher, Education and Instruction
altieri@air.org

Alka Arora

Senior Researcher
aarora@air.org
<https://www.air.org/experts/person/alka-arora>

Lisa Clark

Researcher, Education Statistics
lclark@air.org

Will Tad Johnston

Senior TA Consultant
wjohnston@air.org
<https://www.air.org/experts/person/tad-johnston>

Amanda Mickus

Researcher, Education Statistics
amickus@air.org

Maria Stephens (facilitator)

Senior Researcher
mstephens@air.org
<https://www.air.org/experts/person/maria-stephens>

Reading

Bitnara Jasmine Park

Senior Researcher, Education Statistics
bpark@air.org
<https://www.air.org/experts/person/bitnara-jasmine-park>

Terry Salinger

Institute Fellow, Education and Instruction
tsalinger@air.org
<https://www.air.org/experts/person/terry-salinger>

Yemurai Tsokodayi (facilitator)

Senior Researcher, Education Statistics
ytsokodayi@air.org

Both Subjects

Tessa Grams-Habedank (organization and support)
Research Associate, Education Statistics
tgrams-habedank@air.org

Appendix B: Guide Sheets for Expert Panel Review Meetings

Exhibit B1. Potential Variations to Consider in Comparing PISA and NAEP Mathematics Assessments

Item (or passage) features	Description
Item format	Manner in which common item formats are implemented—e.g., number of choices, features of short v. extended constructed-response items—as well as variation in functionalities (e.g., hot text, drag and drop)
Scoring of constructed-response items	How students earn credit on the items for which they are required to construct a response rather than select a response, including whether scoring is dichotomous (correct/incorrect) or polytomous (includes partial credit level(s))
Clustering and scaffolding	How items with common stimulus material are grouped (clustering) and how those items, and potentially students' responses, are built on one another sequentially (scaffolding)
Mathematizing	Degree to which students must translate a real-world problem situation into a mathematical problem
Role of context	Extent to which interpretation and use of the item's context, or problem situation, is needed for providing a correct answer
Reading load	Amount of reading required by the stimulus material and item
Content specificities	Degree to which specific content varies between assessments
Use and types of representations	Extent and use of visuals that are not strictly text (e.g., tables, graphs, illustrations) in items' stimulus material and the complexity of those visuals

Exhibit B2. Potential Variations to Consider in Comparing PISA and NAEP Reading Assessments

Item (or passage) features	Description
Item format	Manner in which common item formats are implemented—e.g., number of choices, features of short vs. extended constructed-response items—as well as variation in functionalities (e.g., hot text, drag and drop)
Scoring of constructed-response items	How students earn credit on the items for which they are required to construct a response rather than select a response, including whether scoring is dichotomous (correct/incorrect) or polytomous (includes partial credit level(s))
Clustering and scaffolding	How items with common stimulus material are grouped (clustering) and how those items, and potentially students' responses, are built on one another sequentially (scaffolding)
Length and complexity of text	Amount of text and degree of, for example, syntactic and semantic complexity
Authenticity	Degree to which text has been kept in its original form versus manipulated or edited
Explicitness, transparency, and clarity of text structure	How apparent and useful are features of text structure (such as paragraphs or headings)
Familiarity with structure and genre	Degree to which reader is likely to be familiar with the text's structure and genre
Number of features and conditions	Number of pieces of information that need to be located or considered in the text in order to answer the question
Proximity of pieces of required information	How close together in the text are the pieces of information that need to be connected in order to answer the question
Prominence of necessary textual information	How prominent is the information that is needed to answer the question (even if not sufficient by itself to answer the question)
Semantic match between text and task	Degree to which there is a semantic match between the wording of the task and the information that is needed to answer the question (even if not sufficient by itself to answer the question)
Amount of inference required	Degree to which the reader is required to infer non-explicit information from the text
Type of interpretation	How a reader makes meaning from something that is not stated (e.g., making comparisons, finding contrasts, summarizing main ideas)
Nature of knowledge reader needs to bring	Degree to which reader is required to bring outside knowledge to answer the question
Depth of understanding	Depth of understanding of the text required to answer the question