Organisation for Economic Co-operation and Development

EDU/PISA/GB(2018)22

For Official Use **English text only**

22 October 2018

DIRECTORATE FOR EDUCATION AND SKILLS
PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT

Governing Board

PISA SPECIAL EDUCATION NEEDS FEASIBILITY STUDY

46th meeting of the PISA Governing Board

5-7 November 2018
Prague, Czech Republic

Andreas Schleicher, Director for Education and Skills and Special Advisor on Education Policy to OECD's Secretary-General (andreas.schleicher@oecd.org)

JT03437982

Introduction

1. This document provides an overview of the special education needs feasibility study and its results to understand how technology could be used to increase inclusiveness of the Programme for International Student Assessment (PISA) and identify longer-term goals for PISA. The current long term-strategy for PISA covers the period from 2015 to 2024 [EDU/PISA/GB(2013)14] and stresses the importance of making PISA more inclusive. One of the six objectives for the longer-term development of PISA is "to seek ways to widen access of PISA for students with disabilities and other special education needs".

2. At the 44th meeting of the PISA Governing Board in November 2017, the Secretariat presented the project of a feasibility study regarding students with special education needs. The purpose of the feasibility study is i) to investigate how technology can be used to make PISA more inclusive to students with disabilities and other special education needs, and ii) to evaluate the ability of test takers with special needs to access test content using their preferred assistive technology and respond in a digital environment rather than to measure their proficiency in the content being assessed. The results of this study are presented in this report.

3. The PGB is invited to **NOTE** and **COMMENT** on the report.

Background

4. In many countries, children with disabilities represent a significant proportion of students in schools. For example, in the United States, 13% of students receive special education services for a disability (National Center for Education Outcomes, 2018). A large portion of these students also receive testing accommodations such as braille for paper-based tests and screen magnification for computer-based tests. The purpose of this study was to investigate how technology could be used to increase inclusiveness of the Programme for International Student Assessment (PISA) and identify longer-term goals for PISA. Currently PISA is not programmed to accommodate students who require assistive technology to access and interact with computer-delivered assessments. The PISA 2018 Terms of Reference and the longer term goals for PISA, outlined in the position paper by the PISA Governing Board, stressed the importance of making PISA more inclusive (PISA Governing Board, 2013). More specifically, Objective 6 in the position paper was "to seek ways to widen access of PISA for students with disabilities and other special education needs" (PISA Governing Board, 2013: 15). The PISA Governing Board recommended that this could be accomplished "by investigating how technology could be used to increase inclusiveness and then conducting small-scale trials, perhaps embedded in a field trial administration, to determine how various accommodations could be made available and what impact their availability would have on exclusion rates" (PISA Governing Board, 2013: 16).

5. For PISA 2018, Educational Testing Service (ETS) proposed to address this issue by investigating the delivery of a set of items from PISA 2015 on a platform aligned with the World Wide Web Consortium (W3C) accessibility standards, most notably the *Web Content Accessibility Guidelines* (WCAG; W3C, 2008). These standards were developed with international collaboration and with the goal of providing a "single shared standard for web content accessibility that meets the needs of individuals, organisations,

and governments internationally" (W3C, 2005). The guidelines include four main principles that all content must be 1) perceivable; 2) operable; 3) understandable; and 4) robust. In addition, the W3C provides tutorials and authoring tool guidelines on how to author accessible content for many of the item types on PISA. For example, radio buttons, checkboxes and open-ended text boxes have complete tutorials and authoring guidelines because they are frequently used in a wide variety of web content (e.g. government forms, online shopping and online surveys). Even when these standards are followed, however, the availability and variety of assistive technologies will vary across countries and language groups. In addition, some of the item types on PISA do not yet have consistent accessibility standards (e.g. drag and drop). Finally, not all content in PISA items can be easily retrofitted to comply with these guidelines.

6. For these reasons, a small feasibility study was conducted to inform PISA Governing Board decisions on to the feasibility of expanding the inclusion of individuals with disabilities in future assessments from the assessment point of view.

7. The two main goals of this study were to 1) identify which item types can be made accessible[1] to individuals with disabilities using assistive technology and 2) identify which assistive technologies can be supported across different countries and what additional human support might be required. To investigate these questions we focused primarily on the usability of the computer-based testing platform by examining three main areas 1) familiarity with the computer-based testing formats; 2) independence (with or without assistive technology) to access and respond to the test questions; and 3) understanding (both of the science content and the specific task instructions). In addition, we gathered information on the assistive technologies used by students in our sample. It is important to note that the purpose of the study was not to examine aspects of science performance and therefore, no effort was made to score or analyse the students' responses to the items.

[1] Accessibility was defined as adapting the delivery system to be W3C WCAG AA compliant. AA refers to the second level of conformance to the WCAG guidelines. This indicates that the web content meets the minimum level of conformance (level A) and the level AA or an equivalent alternate.

Method

Participants

8. Five PISA member countries[2] (Canada, Dubai [United Arab Emirates], the Netherlands, Scotland [United Kingdom], and Spain) volunteered to participate in this study. The target sample for this study was 50 students with disabilities from all participating countries (10 students from each of the five participating countries). Each country was asked to identify an assistive technology expert to recruit students within the country who met the following five study eligibility criteria:

1) The students were 15 to 18 years old, with a preference for students closer to 15 years old.
2) The students used an assistive technology commonly used in their country to access electronic material.
3) The students had one of four primary disabilities. Target samples by disability group were blindness (*n* =2), low vision (*n* = 3), learning disability (*n* = 3), or physical disability (*n* = 2). These disability groups were selected because the testing accommodations typically used by students in these groups require the use of assistive technologies.
4) The students did not have a cognitive or intellectual disability that would affect their ability to complete the PISA items.
5) The students had the opportunity to learn the science covered in PISA assessments.

9. Table 1 includes the target sample, the achieved sample and the age range of the achieved sample by country and disability group. There were 37 students with usable data across the countries and disability groups. One country exceeded the target sample, one country achieved the target sample and three countries were not able to recruit the number of students requested for the target sample. Recruitment was the responsibility of each of the within-country assistive technology experts, and they had varying levels of access to students. Some countries used their direct connections with mainstream or private schools for students with specific disabilities. Other approaches included using their colleagues with connections to eligible students and working with educational leadership and professional organisations such as SEN (Special Education Needs) technology teams and teachers and regional student support services groups. Recruitment challenges identified by assistive technology experts were school holidays and testing schedules, assistive technology expert availability and an inability to find students who met the diagnostic criteria in the needed timeframe. One student completed most of the study but was not included in the data analysis because it was determined he had multiple disabilities that disqualified him based on the sample criteria. A description of the complete sample by age, gender, disability group and assistive technologies used are described below.

[2] For simplicity, we have used the terms “country” and “countries” to represent participants. The samples, however, were not intended to be nationally representative.

Age

10. Participants ranged in age from 14 to 19 years old, with the majority of students from 15 to 17 years of age. Because the focus of this study was on usability, and recruitment was expected to be a challenge (particularly for low-incidence disabilities), countries were instructed to include older students (16–18 years) if they could not find 15-year-old students who met the criteria because it was expected the older students would interact with the content similarly to 15-year-old students. One participant, however, was 19 years old but enrolled in grade 11. Another student from Scotland was 14 years old and enrolled in S4, which is equivalent to grade 9 in US schools ("Conversion Table", 2018).

Gender

11. The sample included a mix of boys (*n* = 20) and girls (*n* = 17). All country samples included both boys and girls; it should be noted, however, that the students who were blind were overrepresented by girls (7 of 8 students) and the students with learning disabilities were overrepresented by boys (10 of 11 students). The students with low vision and physical disabilities had relatively equal representation of boys and girls.

Disability group definitions

12. Because the diagnostic criteria for some disabilities varied by country, we asked representatives in each country to determine whether there were any significant differences between the US definitions of disabilities included in Annex A and the definitions used in their country. All countries used the US-based definitions to identify their sample. Dubai and the Netherlands indicated they have country – and citywide disability definitions and that there were no major differences in the definitions that the United States uses and those used in their locations. Canada, Scotland and Spain indicated that they did not have countrywide definitions for disability groups.

Assistive technology

13. Assistive technology experts were encouraged to recruit students who were familiar with the most widely used assistive technology (by disability group) in their country. For example, students who are blind or low vision in the Netherlands are more likely to use the SuperNova screen reader and screen magnification software, so we asked that students familiar with SuperNova be recruited. This resulted in a wide use of assistive technology and browsers across countries. Table 2 includes the most common types of assistive technologies the students reported using by country. In some cases, students reported using multiple assistive technologies, and in a few cases, they used no assistive technologies, so the totals do not match the sample sizes for each country. In addition to the three most commonly reported assistive technology categories included in Table 2 (screen reader, text-to-speech, screen magnifier), there were several other types of assistive technology used by students in the sample. These include braille displays, voice recognition (e.g. Dragon NaturallySpeaking), eye-gaze and eye-tracking software, and switch access. Switch access is an assistive technology used by individuals with limited physical abilities or cognitive impairments that allows users to use a switch (button) in place of a keyboard, mouse, or touchscreen. Assistive technology experts were asked to collect the level of experience each student had with the assistive technology that they used. The majority of students (59%, *n* = 22) were reported as advanced, near a quarter, (22%, *n* = 8) were reported as intermediate, one student (3% of the sample) was reported as beginner and 16% (*n* = 6) of the sample were not rated because they did not regularly use assistive technology.

All 11 students from the Netherlands were advanced AT users. All of the students who did not report using assistive technology were from Spain. This may have accounted for some country-level differences reported below in the results section of this report.

Materials

Background information questionnaire

14. A set of questions was prepared for students to complete. The questionnaire could be administered at the start or the end of the study session. The questions included the student's age, sex, current grade, disability information, types of assistive technology use, level of experience with assistive technology, type of computer typically used, primary operating system used and post-high school plans.

Test items

15. The selection of items in this feasibility study was completed by a team of ETS staff with expertise in accessibility of assessments, W3C Web Content Accessibility Guidelines and PISA development. Items from recently released PISA 2015 science assessments were reviewed and considered for selection. PISA 2015 science items were selected because these items contain numerical components (similar to PISA mathematics items) and text structure components (similar to PISA reading items). Released PISA items were used in this study because the test items would not have to be carefully protected to maintain their confidentiality status, and the existing national translations ensured the items' availability in all the necessary languages.

16. While cognitive labs can provide greater insights into accessibility challenges and possible solutions that students encounter during testing, the one-on-one administration limits the number of items that can be included. For this reason, the focus of item selection was on exposing students to a wide variety of item type (with varying levels of accessibility challenges). Based on these criteria, a decision was made to include the general PISA orientation and one science task. The one science task was the *Energy-Efficient House* task[3], which, despite the fact it is one of the most complex released units, contains a variety of scenario-based item types that are included in the PISA science assessments (as well as the PISA mathematics and the PISA reading assessments). The full set of items included examples of basic navigation tasks (move to the next item), eight distinct item types and combined tasks (integration of two complex items types in a single task). See Annex B for screenshots of each of the following interactions:

1) Basic navigation
2) Click on a choice (also referred to as a radio button or multiple choice)
3) Click on choices in a table (select answers among rows of choices in a table)
4) Click on one or more boxes (also referred to as a checkbox or multiple selection multiple choice)
5) Type text (also referred to as type your answer or open-ended text entry)
6) Enter a number (open-ended numeric entry)

[3] https://www.oecd.org/pisa/test/PISA2015-Released-FT-Cognitive-Items.pdf

7) Drop-down menus (select one answer per list of choices in a menu)

8) Drag and drop (select an answer by moving an option from one column to a corresponding answer in another column)

9) Run a simulation (requires students to run a simulation to populate a data table)

10) Combined task (drag and drop with a simulation)

Protocol

17. A protocol was used by the assistive technology experts serving as the interviewers in the cognitive lab sessions to facilitate each study session with the student research participants. The protocol provided directions to help the interviewer guide the student through the session, including specific information and instructions to convey to the student. It also included directions to the interviewer to record specific information as the student worked on the study tasks and provided questions to ask the student after the completion of each study task. The protocol included direct questions and sections to record observations regarding student familiarity, independence and understanding. The observation section included concurrent notes tables and mostly discrete categories that allowed for the interviewer to easily capture information by selecting the appropriate response to each test item or set of interactions on a screen. Assistive technology experts were asked to include details on what they observed (student actions and verbalisations) to justify each of the discrete categories. Before moving to the next screen of information, the interviewer asked the student retrospective questions to obtain more information about the aspects of the item that work well and to identify the specific challenges the student encountered. In practice, however, these questions yielded little additional information, because most of the assistive technology experts chose to include detailed feedback in prior open-ended questions to justify their ratings.

18. The protocols were provided to the assistive technology experts in English as a Word document, and electronically in English as a Google survey. Assistive technology experts were expected to be familiar enough with the protocol to have a good understanding of the topics to observe and types of questions to ask (see discussion of training and calibration of ratings in following section on procedures). They were familiar enough with the purpose of each question and section of the protocol that they could reword the questions or directions to the students if the original protocol language was not understandable. Assistive technology experts translated the protocol to their country's home language and communicated with the participants in each student's choice of language. This sometimes included a combination of English and a second language.

W3C test delivery platform

19. The PISA test uses standard HTML and web browsers for delivery. The orientation and science task were delivered via the Internet using individual logins and passwords. The content was enhanced and tested to ensure that it would display using standard web browsers (Firefox, Chrome, Safari and Internet Explorer). It was strongly recommended that participants use the latest versions of these browsers. For Internet Explorer, only the latest version (11.0) was supported.

20. Once the item types and PISA task were selected, each screen was evaluated by a team with expertise in the accessibility and presentation of test content for individuals with disabilities. The programmers and accessibility experts worked iteratively to improve the experience for the study participants so that they could access and interact with each

component of the test in a meaningful way. Ultimately, the test content delivered in this study was compliant with WCAG 2.0. The major changes that were made include:

1) Making sure each frame or region[4] was able to be accessed
2) Hiding frames or regions with redundant information or disabled aspects
3) Adding keyboard accessibility
4) Adding Accessible Rich Internet Application (ARIA) labels to improve interoperability with assistive technologies
5) Preparing functional versions of item types for which there is no standard for the best format
6) Preparing alternative text, long descriptions and off-screen instructions for users of screen readers
7) Preparing paper-based instructions for administration if assistive technology failed to work with the content

21. There were some limitations to the changes that could be made in the timeframe and budget of this effort. Most notably, there was not sufficient time in the study to create translations for new content resulting from the accessibility review. This included any wording changes to the text of the directions (for example, using the mouse-centric language, “click one or more boxes”) or for newly created alternative text for images or to describe tables. See Annex C for a more comprehensive description of the changes that were made to the programming of the content. Additionally, the general organisation of the display and the screen layout was not changed from the original versions of the units.

Browsers

22. The test required a screen resolution of 1024 x 768 or greater. All four browsers that were tested and one additional browser, Microsoft Edge, were used to access the study content. Table 3 shows the browsers that were used in this study.

Procedure

23. To investigate our two main research questions, we used a qualitative research design. The cognitive lab interview served as the primary source of data to understand the item types and the usability of the web-based items with the students’ assistive technology. This approach provides a way to collect information to evaluate comprehension and understanding processes (Leighton, 2017). In this study, students were not instructed to think aloud as they sometimes may be in cognitive lab interviews. This action would be inappropriate, and, in fact, counterproductive, for the students who were using assistive technology with audio output including screen readers and text-to-speech software. The students would be less able to concentrate and comprehend the content if they were required to speak their thoughts aloud at the same

[4] Web pages are typically organised by region or section for the different functionalities and content areas on the page. Web pages usually have regions for the page header, page footer, navigation, main content and may include others depending on the content and purpose of the web page. A well designed web page uses regions to organise each functional section which allows assistive technology users to easily find and navigate by region.

time. In addition, Branch (2000) indicated that some students have difficulty with problem solving and talking out loud simultaneously.

24. In addition to the interviewer's observations of student interactions with the PISA content, the procedure in this study was designed to capture feedback from the students immediately after they completed a task after one screen of information. We collected what Ericsson and Simon (1993) described as Type-3 verbalisations, explanations of what the students were thinking elicited through retrospective questions asked by the interviewer. In addition to the scripted retrospective questions included in the protocol, assistive technology experts were able to ask nonleading clarifying questions based on what they observed in close proximity to when the instance happened. For example, an assistive technology expert may ask, "I saw you click the mouse three time when you were selecting the red option in the simulation. Can you tell me more about what was happening?" The retrospective questions were designed to capture whether the students knew what they were supposed to do based on the wording of the instructions and items even if their assistive technology was not allowing them to interact with the content as they wished.

Selection of assistive technology experts

25. The assistive technology experts played a critical role in this study to recruit the participants and collect and enter the data for this study. These individuals were selected by the participating countries using the following criteria: 1) have some experience working with 15-year-old students who use assistive technology with the following disabilities: blindness, low vision, learning disabilities, and physical disabilities; 2) have direct access to 15-year-old students who use assistive technology with the following disabilities: blindness, low vision, learning disabilities, and physical disabilities; and 3) have knowledge about the assistive technology used by 15-year-old students with print and physical disabilities while accessing electronic materials and have the ability to obtain comprehensive information about the assistive technology they have little or no knowledge about.

Training of assistive technology experts

26. The primary responsibility of each assistive technology expert was to recruit and collect data from 10 students in their country successfully. In order to complete this work accurately, each assistive technology expert was expected to participate in a one-day in-person training to review the study materials, including the PISA content to be administered. The one-day training occurred in Princeton, New Jersey, USA. Two of the assistive technology experts were able to attend in person. Representatives from two of the other countries participated via the Internet, and one of the assistive technology experts completed the training remotely on a separate date.

27. The training included an overview of the functionalities for each screen and the modifications that were made to improve the accessibility of the content were reviewed. The assistive technology experts were also provided guidance on how to recruit students from their countries. The majority of the training focused on detailed instructions on how to use the protocol to collect and enter their data. This included watching videos of researchers answering test items with assistive technologies typically used by students with disabilities in the target sample. The videos were developed to demonstrate different levels of understanding and independence through both actions and verbalisation. The assistive technology experts used the protocol to rate the video samples and then

received feedback from the research team. In addition, each assistive technology expert completed an administration (either with each other during the in-person training or with another adult for the off-site training). Complete data (ratings and justification for the ratings) was reviewed by a researcher prior to the start of data collection to ensure that all assistive technology experts had common understanding of the discrete ratings and had provided details to allow for researchers to review responses.

Data collection

28. The study sessions occurred at the students' schools in locations that were conducive for a student to take a web-based computer-delivered test. The needed assistive technology for the student was made available. Each study session began with the assistive technology expert providing instructions to the student. The assistive technology expert thanked the student for their willingness to participate in the study, indicated that participation was voluntary, they could withdraw at any time without penalty, their anonymity would be protected and data captured and recorded in this study would be stored securely and would not be used as a measure of science performance. The steps of the session were explained, and the student was reminded that the purpose of the study is to learn and gather information on what works and does not work with the formatting of the materials and that the purpose is not to measure individual student science skills.

29. The students began with the PISA orientation. The student worked independently to read, interact with and answer whatever was on each screen of information. During this time, the assistive technology expert recorded observations on the student's familiarity with, independent interaction with and understanding of what was presented on that particular screen. The protocol also required the assistive technology expert to provide evidence (behaviours or verbalisation by the student) to support their ratings. If the student had a question or got stuck, the assistive technology expert was instructed to document the question or problem and assist the student to allow them to move forward. Once the student completed each screen, the assistive technology expert asked the retrospective questions before instructing the student to move to the next screen. This process was repeated for each of the orientation screens and then for each of the Energy Efficient House screens.

Data entry

30. All data captured from the protocol was entered directly into a Google survey or a Microsoft Excel file in English. This data included the assistive technology expert's observations and notes of student responses to their questions. It also included the assistive technology expert's ratings on the student level of independence and the student self-report on familiarity and understanding (for each item type and interaction).

Results

31. To answer our primary research question (i.e. identify which item types can be made accessible to individuals with disabilities using assistive technology) we examined each item type (or interaction) across three areas 1) familiarity[5]; 2) independence; and 3) understanding. Each of these three areas are then summarised across all item types and interactions. To facilitate interpretation of the data for decision making, each item type was classified into one of three categories (accessible, partially accessible and not yet accessible). Accessible items were familiar to 80% or more of the study participants and group differences[6] were not found. Partially accessible item types fell into two groups: i) familiar to = 80% (or more) of students but with group differences identified or ii) familiar to 60% to 79% of the students tested and no differences were noted between countries or disability groups. For example an item that was familiar to 90% of students overall but was familiar to only half of students with physical disabilities was rated as partially accessible). Finally not yet accessible items were familiar to less than 60% of students in our sample. We chose the language not yet accessible because future W3C Web Accessibility Guidelines (WCAG) are likely to include specifications for some of these item types and over time students will become more familiar with them. While the above classification criteria were described using the example of familiarity, we used the same criteria for the independence and understanding categories. In some cases the assistive technology experts did not provide complete data for all item types, so researchers evaluated the open-ended responses to determine the student's level of familiarity, independence and understanding. When this was not possible the response was rated as "other". Open-ended responses were reviewed to provide greater understanding of the ratings and are included for illustrative purposes below.

Familiarity

32. Student familiarity with the computer-based testing item types and interactions was assessed by asking students, "Have you ever seen information like this on a computer before?" Answer options included "Yes", "Yes, but not on a test", "No" and "Not sure or other". Most students (86%) responded that they had seen click on choice items (i.e. single-selection multiple-choice items represented by radio buttons) before on a computer-based test, resulting in an accessible classification. However, the majority of item types were classified as partially accessible, and three item types were classified as not yet accessible (i.e. click on more than one box, drag and drop, run a simulation). For three of the partially accessible item types, the student's familiarity varied between countries. In these cases, students in Spain reported being less familiar with

[5] While familiarity is not a prerequisite for accessibility, we have included it in the overall rating of accessibility because student familiarity with item types and interactions can impact both independence and understanding.

[6] Given the small sample sizes of some groups, a rough criterion for group differences was selected. If 50% (or more) of students from each country or disability group were not familiar with the item type, it was classified as partially accessible or not yet accessible. This same criterion was used for student independence and understanding. In addition, we did not include Canada or Dubai in country-level difference because the differences could not be meaningfully compared due to the very small sample sizes for those countries.

computer-based testing than the students selected for participation in Finland and Scotland. Half (five of ten) of the students in Spain reported that they were not familiar with the navigation features of computer-based testing and two of the items types (type text and enter number). The majority of students in the Netherlands (nine of 11) and Scotland (eight of nine) were familiar with this type of navigation on a computer-based test. On closer examination, it seemed that these differences were primarily due to four of the five students with visual disabilities (blind and low vision) in Spain reporting that they had not seen this type of information on a computer before. For the type text and numeric entry items, we also noted that half of the students (four of eight) who had not seen the text type item on a computer-based test before were from Spain.

33. Two of the three items types rated as not yet accessible for familiarity also showed difference by disability subgroups. These were the item types that used drag and drop and simulation. For the drag-and-drop item type, none of the blind students had encountered this type of question on a computer-based test and only two of the 11 low vision students had. The majority of students with physical (six of seven) and learning (nine of 11) disabilities reported being familiar with this item type. Findings were similar for the simulation item type, where more than half of students with learning (six of 11) and physical (four of seven) disabilities reported seeing this type of test item before, but less than half of students with who are blind (three of eight) and with low vision (two of 11) reported encountering this type of question on a test in the past.

34. It should be noted that student's familiarity did not appear to be related to levels of independence or understanding during testing. For relatively straightforward tasks like navigating to the next item, the prior familiarity with computer-based testing did not impact either independence or understanding as described in the next two sections.

Independence

35. Student independence with the computer-based testing item types and interactions was assessed by asking, "Could the student independently access everything on the screen?" Each screen included one interaction (as described in the prior section on familiarity). However, students were also asked to engage with a complex screen that included multiple interactions (i.e., drag and drop within a simulation). Answer options included "Yes", "Partially", "No" and "Not sure or other". In two cases the students experienced technological difficulties, and the assistive technology expert was unable to load the screen. These cases were classified as not sure or other because it was unclear if those technological difficulties were due to interactions between the assistive technology and the testing platform or issues with school bandwidth. To facilitate interpretation of the data for decision making, each item type was similarly classified into one of three categories (accessible, partially accessible and not yet accessible) using the same criteria used previously for familiarity. In addition, the assistive technology expert asked some follow up questions on the types of challenges students encountered and possible solutions.

36. All item types and interactions were rated as accessible (three item types) or partially accessible (6 item types and basic navigation) for student's level of independence. The three accessible item types were click on one or more boxes, click on choice in table and type text. One surprising finding was that the two simplest interactions (basic navigation and click on a choice [radio button]) were not accessible due primarily to problems encountered by students who are blind or who have low vision. For the radio button item type, half of students who are blind (four or eight) were only partially proficient the first time they encountered this interaction (general introduction). This was primarily

due to the screen reader not reading all portions of the text. One common report was, “JAWS did not read the text in yellow on the left-hand side of the screen.” This text included directions for the item type. In addition, two students were able to enter an answer but did not know how to change the answer. For the basic navigation interaction, most challenges were experienced by students who are blind (four of eight). Assistive technology experts reported that students could not initially find the “next” button. Further, in some cases the student placed the mouse (or tabbed to) the correct area of the screen, but the hover text did not appear or was not read aloud by the screen reader. In these cases, the assistive technology experts reported that they had to provide additional verbal instructions to the student and that on subsequent screens the student was independent. In addition, there were some differences by country for this interaction (i.e. in Spain, four of the ten students needed additional verbal instruction from the assistive technology expert before they understood what to do, while in the Netherlands, none of the students required additional instructions).

37. Five other item types (enter a number, drop-down, drag and drop, run a simulation and complex interaction) were also rated as partially accessible primarily (but not exclusively) due to challenges experienced by students who were blind or had low vision. Half of students who were blind could not independently access all parts of each of these item types. This was most pronounced on the simulation item type: All of the students with learning and physical disabilities were able to independently access the item, but only three of eight students who were blind and six of ten students with low vision were able to do so.

38. One challenge that frequently interfered with students’ independence (but not specific to any one item type) was related to challenges of the layout of the screen. Over the course of the administration the assistive technology experts noted the layout as a challenge for nine of the 37 students. For example, one assistive technology expert reported that for a low vision student, “The complexity of screen layout and instructions, and the way in which instructions refer to ‘other areas’ of the screen which are not visible to the student made her confused about how to interact with the screen.” Another assistive technology expert reported similar issues for a student with cerebral palsy (physical disability): “The main challenges for the student were 1) the complexity of the layout on screen and the number of separate areas and when magnified 2) confusion over the wording of instructions.” In these types of cases, the assistive technology experts generally reported that they had to offer additional verbal explanation or instructions to the students to move on. It should be noted that the steps to alter the organisation of the content and to rewrite the orientation and instructions were outside the scope of this research effort to allow for existing PISA items and translations to be used.

Understanding

39. Student understanding of how to answer each item type and interaction was assessed by asking, “Did the student understand the functionalities and content presented on this screen?” Assistive technology experts asked students directly and also prompted further when the student seemed confused or unsure. Each screen included one interaction (as describe in the prior section) and one screen with a complex interaction (i.e. drag and drop within a simulation). Answer options included “Yes”, “Partially”, “No” and “Not sure or other”. As was described earlier, two cases where students experienced technological difficulties were classified as not sure or other because it was unclear if those technological difficulties were due to interactions between the assistive technology and the testing platform or issues with school bandwidth. To facilitate interpretation

of the data for decision making, each item type was similarly classified into one of three categories (accessible, partially accessible and not yet accessible) using the same criteria used previously for familiarity and independence. In addition, the assistive technology expert asked some follow-up questions on the types of challenges students encountered and possible solutions.

40. Overall, assistive technology experts reported that students understood the functionality for basic navigation and six of the item types (click on choice [radio button], click on one or more boxes [checkbox], click on choice in table, type text, type number and drop-down menus). The remaining item types were rated as partially accessible for various reasons. It should be noted that the drop-down menu item, while rated as accessible for student understanding, prompted several suggestions for improvement, and one student with a learning disability only understood after the assistive technology expert read it aloud. One common suggestion (made by five students) was to adjust mouse-centric language (e.g. not using phrases such as "hold the mouse button" or "release the mouse button") but did not suggest specific alternatives.

41. Despite the fact that most students reported to the assistive technology expert that they understood what they were supposed to do, several assistive technology experts commented that they were not entirely convinced that the student could complete the necessary actions on an actual assessment. This was particularly true for the click-on choices in a table item for which one assistive technology expert stated, "The student said that she could not work out how to navigate the rows and columns in the table." Another assistive technology expert reported, "I think the student understood that he was expected to make selections in the table, but I am not convinced that he would be able to do so in an actual assessment" and "The student spends more time than usual filling in the circles. Asked about it, he states that it has taken him a while to find them but he does not suggest any improvement."

Additional observations

42. Although not a purpose of this study, students did show understanding of the science content knowledge assessed on the test items. For the most complex item type that included a drop-down and text entry, students were asked, "Explain the differences in energy consumption by describing what happens to solar radiation when it hits these two different colours of roof?" Fifty-nine percent of the students (n = 22) answered correctly. It should also be noted that despite the item referencing colour, six of the eight students who are blind were able to answer correctly.

43. The small sample size and lack of consistency across assistive technologies made it difficult to answer our second research question (identify which assistive technologies can be supported across different countries and what additional human support might be required). However, the screen reader technologies (critical for students who are blind) presented the most problems across all items types. Assistive technology experts reported that they needed to read aloud content on the screen and in some cases make selections using a mouse for students to continue.

Summary and Recommendations

44. Overall there were very few items types that were classified as Accessible on all three criteria (familiarity, independence and understanding). See Table 7 for a summary across all three categories. The majority of students in our sample, however, understood what they were supposed to do and were able to answer test items independently. This was less true for students who were blind or had low vision and relied on screen readers, and several assistive technology experts noted problems with JAWS in Internet Explorer.

45. The current PISA platform is a legacy system that has been developed over many years. While the platform can accommodate the W3C accessibility guidelines, retrofitting existing assessment units will require additional work. For example, numerous lines of code were needed to allow screen readers to navigate to and identify buttons. It was outside of the scope of this research effort to alter the platform or test content and directions. Development of alternative formats for existing items and the overall display of the PISA units would increase the compatibility of this content with the W3C guidelines and with common assistive technologies. This development would entail additional effort on the part of the PISA test development and platform teams, as well as participating countries, with associated impacts on both scope of work and budget.

46. Below are three recommendations for the PISA Governing Board to consider as they evaluate the results of this study and next steps in achieving Objective 6, "To seek ways to widen access of PISA for students with disabilities and other special education needs".

Develop accessible test forms

47. Most of the item types were at least partially accessible and could be completed with high levels of independence and understanding by the majority of students in our sample. One initial step in moving towards accessible assessments is to develop a test form that includes a limited pool of item types. Based on the results of this study, a recommendation would be to include less complex items types (i.e. click on choice, click on one or more boxes, click on choice in table and type text) which were rated as Accessible and Partially Accessible with few differences across disability groups. This should not present a significant limitation, as these item types constitute the majority of the PISA item pool. It would also be advised to consider altering the layout of this test form to avoid the challenges encountered with navigation through multiple sections of the screen. One format used for braille and screen reader users in the Smarter Balanced Assessment is termed streamline view and organises all content of an item in a linear fashion from the top to the bottom of the screen. "This designated support provides a streamlined interface of the test in an alternate, simplified format in which the items are displayed below the stimuli" (Smarter Balanced, 2018a: 11). Another way to increase access to students would be to revise the directions and evaluate the language used in the general orientation with the goal of providing appropriate and useful instructions for nonsighted students and students who do not use a mouse. It should be acknowledged that the time and efforts to complete these recommendations could be significant.

Develop guidelines

48. In cases where students could not independently access items, the assistive technology experts reported reading aloud test content or moving the mouse to advance to the next screen or make a selection for the student. These approaches are common accommodations for students with disabilities. To improve the accessibility and standardised delivery, it is recommended that a guidelines for human assistance be developed. These guidelines provide instructions for human assistance on how to record students answers and how to read aloud test content. See Smarter Balanced, 2018b; Educational Testing Service, 2002a; and Educational Testing Service, 2002b for examples of scribe and human reader guidelines. This approach would impact the PISA procedures and guidelines at the international level and would require additional country-level resources during survey operation.

Incorporate accessibility training

49. A final recommendation is to incorporate training on accessibility into future item development efforts. This recommendation is based partially on the preparation of test items for the study. Most items considered for inclusion in this study require retrofitting for accessibility (see Annex C for the complete list). The main adjustments were to address two principles (perceivable and understandable) of Web Content Accessibility Guidelines (W3C). For example, the choice of colours in the Energy Efficient House task relied on vision (or alternate text descriptions) to make the task perceivable and understandable. In this study, alternate text was added to describe images and the layout. This type of retrofitting in the future will require additional time and add to the translation and linguistic quality control aspects of the project. Another example that was mentioned by students is the use of mouse-centric language that may not be appropriate for test takers using an alternative pointing device or a screen reader, such as, "click on a choice". Several K12 testing programs in the United States have incorporated training for designers in accessibility by using the principles of universal design for learning (Meyer, Rose and Gordon, 2014). These principles guide designers to consider multiple means of representation, multiple means of action and expression, and multiple means of engagement for each test item. This type of training could reduce the need for retrofitting items.

Limitations

50. This study included some limitations that are important to note.

51. First, the sample size was relatively small and is not representative of all students in a particular country or all potential PISA test takers with special education needs.

52. Second, the number of items and tasks included in a qualitative study is small and cannot generalise to all PISA items. While we did make an attempt to cover the range of interactions, it is possible that some items will include more demands on students with disabilities due to the content of the task (e.g. graphics that are difficult to describe in alternative text) and not the item type.

53. Third, while we attempted to recruit only test takers with assistive technology, it was clear from the reported data that not all students used assistive technology during testing and that some students were not aware of the full potential use of the available assistive technology. For example, one assistive technology expert reported that he

observed one student with low vision (who was using Google Chrome magnification) "peering closely at the screen. I showed him how to use CTRL+ on the keyboard to zoom in – he said he had not come across it before. He said it was useful to zoom in a little to enlarge the text size." Future studies may want to ensure that all of the country level assistive technology experts have access and time to recruit students who are proficient in the assistive technology they use.

54. Finally, while cases describing individual difficulties were provided in the results sections above to illustrate challenges students had, the specific causes of the difficulties were unable to be systematically evaluated in this study which is a limitation. The difficulties students faced were likely due to one of several possible factors that could not be identified with certainty due to the level of detail that the assistive technology experts were able to capture. The difficulties could have been student-based challenges, including a student's lack of mastery in their assistive technology, or lack of exposure to the item type which could limit a student's ability to navigate and make selections based on their unfamiliarity of the structure and expectations of the item types. There could have been software-based difficulties that did not allow a student to have full access or ability to interact with the test content. There could also have been web interface challenges such as missing descriptive information for assistive technology users, like an icon label, or an inefficient presentation of information for students using assistive technology. In order to evaluate causality, the reason for the challenge encountered would have to have been captured for each instance of a difficulty which was not possible during this data collection. Future research efforts could examine these challenges by implementing a few procedures. First a detailed background survey of the students' experience of the assistive technology that they use would need to be collected. Ideally, this qualitative study would be limited to students who have a strong mastery of their assistive technology. This eligibility criteria would increase the time to recruit the intended sample. In addition, this background survey should capture how much experience the students' have with each item type and functionality included in the study. This background survey should be completed by a teacher who is very familiar with the student to ensure this information is accurate. Lastly, the study sessions should be administered by an expert of the particular assistive technology the student is using, and preferably, someone who knows the student, such as the student's teacher or assistive technology aide in their school. These sessions should video record the screen and input activity (for example, standard or refreshable braille keyboard, mouse, alternate pointing device, etc.) so the difficulties may be systematically coded by the possible challenges described above.

Table 1. Targeted Sample and Achieved Sample

Sample size	Target sample per country		Canada		Dubai		Netherlands		Scotland		Spain		Disability group
	N	Age	*N*	Age range	*N*	Age range	*N*	Age range	*N*	Age range	*N*	Age range	
Blind	2	15	1	15	0	N/A	3	16-17	2	15-17	2	15	8 (1 boy, 7 girls)
Low vision	3	15	0	N/A	2	15-16	3	14-19	3	15-18	3	16	11 (6 boys, 5 girls)
Learning disability	3	15	2	17-18	1	17	3	15	2	14-15	3	16-17	11 (10 boys, 1 girl)
Physical disability	2	15	0	N/A	1	18	2	14-16	2	15-17	2	16	7 (3 boys, 4 girls)
Total	**10**		**3**		**4**		**11**		**9**		**10**		**37 (20 boys, 17 girls)**

Table 2. Assistive Technology Used by Accessibility Function Category

Category	Brand	Canada	Dubai	Netherlands	Scotland	Spain
Screen reader	JAWS				2	2
	Nonspecified [1]				1	
	NVDA			1		1
	SuperNova [2]			2		
	VoiceOver				2	
Text-to-speech	ClaroRead			1		
	Kurzweil	1		2		
	Nonspecified [1]	1			3	
Screen magnifier	iPad Apple OS	1			2	
	Nonspecified [1]		1		1	2
	SuperNova [2]			2		
	Windows 10 magnifier			1		
	ZoomText			1		1

Note 1: In these instances, no specific brand or model was reported.
Note 2: Four students indicated they used SuperNova: two used SuperNova as a screen reader, and two used SuperNova for screen magnification.

Table 3. Web Browsers Used by Country

	Canada	Dubai	Netherlands	Scotland	Spain	Browser totals
Chrome	2	3	4	2	9	20
Firefox			1		1	2
Internet Explorer			4	4		8
Microsoft Edge			1			1
Safari	1	1	1	3		6

Table 4. Familiarity with Item Type or Action

Action	Symbol [1]	Percentage	Total	Yes	Yes, but not on a test	No	Not sure/other	Country differences	Disability differences
Basic navigation	≈	70%	37	26	2	8	1	Yes	
Click on choice	✓	86%	37	32	2	1	2		
Click on one or more boxes	X	43%	37	16	8	11	2		
Click on choice in table	≈	78%	37	29	3	4	1		
Type text	≈	76%	37	28	4	4	1	Yes	
Enter number	≈	73%	37	27	3	5	2	Yes	
Drop-down menu	≈	76%	37	28	3	5	1		
Drag and drop	X	46%	37	17	7	11	2		Yes
Run a simulation	X	41%	37	15	4	16	2		Yes

Note: [1]✓ ≥ 80% and no differences by disability group or country, ≈ 60%–79% with no differences by disability or country or ≥ 80% with differences by disability group or country, X < 60%.

Table 5. Ability to Interact Independently with Item Type or Action

Action	Symbol[1]	%	Total	Yes	Partially	No	Not sure/Other	Country differences	Disability differences
Basic navigation	≈	76%	37	28	5	3	1	Yes	Yes
Click on choice	≈	78%	37	29	7	0	1		Yes
Click on one or more boxes	✓	89%	37	33	1	2	1		
Click on choice in table	✓	84%	37	31	3	2	1		
Type text	✓	81%	37	30	5	1	1		
Enter number	≈	76%	37	28	5	3	1		Yes
Drop-down menu	≈	81%	37	30	3	3	1		Yes
Drag and drop	≈	73%	37	27	3	6	1		Yes
Run a simulation	≈	78%	37	29	4	3	1		Yes
Combined item types	≈	70%	37	26	4	7	0		Yes

Note: [1]✓ ≥ 80 and no differences by disability group or country, ≈ 60%–79% or ≥ 80 with difference by disability group or country, X < 60%.

Table 6. Understanding of Functionality of Item Type or Action

Action	Symbol	Percentage	Total	Yes	Partially	No	Not sure/other	Country differences	Disability differences
Basic navigation	✓	97%	37	36	0	0	1		
Click on choice	✓	95%	37	35	0	0	2		
Click on one or more boxes	✓	97%	37	36	0	0	1		
Click on choice in table	✓	92%	37	34	0	1	2		
Type text	✓	86%	37	32	2	2	1		
Enter number	✓	95%	37	35	0	0	2		
Drop-down menu	✓	95%	37	35	1	0	1		
Drag and drop	≈	86%	37	32	2	1	2		Yes
Run a simulation	≈	76%	37	28	0	7	2		
Combined item types	≈	62%	37	23	11	3	0		Yes

Note: [1]✓ ≥ 80 and no differences by disability group or country, ≈ 60%–79% or ≥ 80 with difference by disability group or country, X < 60%.

Table 7. Composite Table of Independence, Familiarity, and Understanding Items Percent Correct

Action	Familiarity symbol[1]	Familiarity percentage (*n*)	Independence symbol[1]	Independence percentage (*n*)	Understanding symbol[1]	Understanding percentage (*n*)	Country differences[2]	Disability differences[2]
Basic navigation	≈	70%	≈	76%	✓	97%	I, F	I
Click on choice	✓	86%	≈	78%	✓	95%		I
Click on one or more boxes	X	43%	✓	89%	✓	97%		
Click on choice in table	≈	78%	✓	84%	✓	92%		
Type text	≈	76%	✓	81%	✓	86%	F	
Enter number	≈	73%	≈	76%	✓	95%	F	I
Drop-down menu	≈	76%	≈	81%	✓	95%		I, U
Drag and drop	X	46%	≈	73%	≈	86%		I, F, U
Run a simulation	X	41%	≈	78%	≈	76%		I, F
Combined item types	NA	NA	≈	70%	≈	62%		I, U

Note: [1]✓ ≥ 80 and no differences by disability group or country, ≈ 60%–79% or ≥ 80 with differences by disability group or country, X < 60%.
[2] F = familiarity, U = understanding, I = independence. When denoted, the item had differential results based on either country or disability.
[3] The total sample size for which the percentages were calculated was n = 37.

References

Branch, J.L. (2000), "Investigating the information-seeking processes of adolescents: The value of using think-alouds and think-afters", *Library and Information Science Research*, Vol. 22, 371–392, Elsevier, Amsterdam, https://doi.org/10.1016/S0740-8188(00)00051-7

Conversion table of USA grades to school years. (2018). Retrieved from the Pearson Clinical website: https://www.pearsonclinical.co.uk/Sitedownloads/Miscpdfs/Gradetoage.pdf

Educational Testing Service (2002a), *ETS guidelines for a test reader*. Retrieved from https://www.ets.org/s/disabilities/doc/guidelines_for_test_reader.doc

Educational Testing Service (2002b), *Guidelines for a scribe*. Retrieved from https://www.ets.org/disabilities/resources/scribe/

Ericsson, K., and H. Simon (1993), *Protocol Analysis: Verbal Reports as Data* (2nd ed.), MIT Press, Boston.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004).

Leighton, J.P. (2017), *Using Think-aloud Interviews and Cognitive Labs in Educational Research*, Oxford University Press, New York.

Meyer, A., D.H. Rose and D. Gordon (2014). *Universal Design for Learning: Theory and Practice,* CAST, Wakefield, MA.

National Center for Education Outcomes (2018), *Students with Disabilities*. https://nceo.info/student_groups/students_with_disabilities

OECD Programme for International Student Assessment (2015), *PISA 2015 Released Field Trial Cognitive Items*, www.oecd.org/pisa/test/PISA2015-Released-FT-Cognitive-Items.pdf

PISA Governing Board (2013, October, 14), *Beyond PISA 2015: A Longer Term Strategy of PISA*, OECD, Paris, www.oecd.org/callsfortenders/ANNEX%20A.pdf

Smarter Balanced Assessment Consortium (2018a), *Usability, Accessibility, and Accommodations Guidelines*, https://portal.smarterbalanced.org/library/en/usability-accessibility-and-accommodations-guidelines.pdf

Smarter Balanced Assessment Consortium (2018b), *Read Aloud Guidelines*, portal.smarterbalanced.org/library/en/read-aloud-guidelines.pdf

W3C (2005). Web *Content Accessibility Guidelines* (WCAG): Overview, www.w3.org/WAI/standards-guidelines/wcag/

W3C (2008, December 11), *Web Content Accessibility Guidelines (WCAG) 2.0*, www.w3.org/TR/2008/REC-WCAG20-20081211/

Annex A. US Definitions[7] of Study Disability Groups

Learning Disability

An individual with a learning disability has a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write, spell or to do mathematical calculations. The term includes such conditions as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia and developmental aphasia. The term does not include learning problems that are primarily the result of visual, hearing or motor disabilities; of intellectual disability; of emotional disturbance; or of environmental, cultural, or economic disadvantage.

Low Vision and Blindness

An individual with low vision has impaired vision that, even with correction, adversely affects a students' educational performance. The distinction between low vision and blindness is that individuals who are considered to have low vision have some partial sight.

Physical Disability

An individual with a physical disability has a severe orthopaedic impairment that adversely affects the student's educational performance. The term includes impairments caused by a congenital anomaly, impairments caused by disease (e.g. poliomyelitis, bone tuberculosis), and impairments from other causes (e.g. cerebral palsy, amputations and fractures or burns that cause contractures).

[7] Given the small sample sizes of some groups, a rough criterion for group differences was selected. If 50% (or more) of students from each country or disability group were not familiar with the item type, it was classified as partially accessible or not yet accessible. This same criterion was used for student independence and understanding. In addition, we did not include Canada or Dubai in country-level difference because the differences could not be meaningfully compared due to the very small sample sizes for those countries.

Annex B.

Figure B.1. Basic Navigation

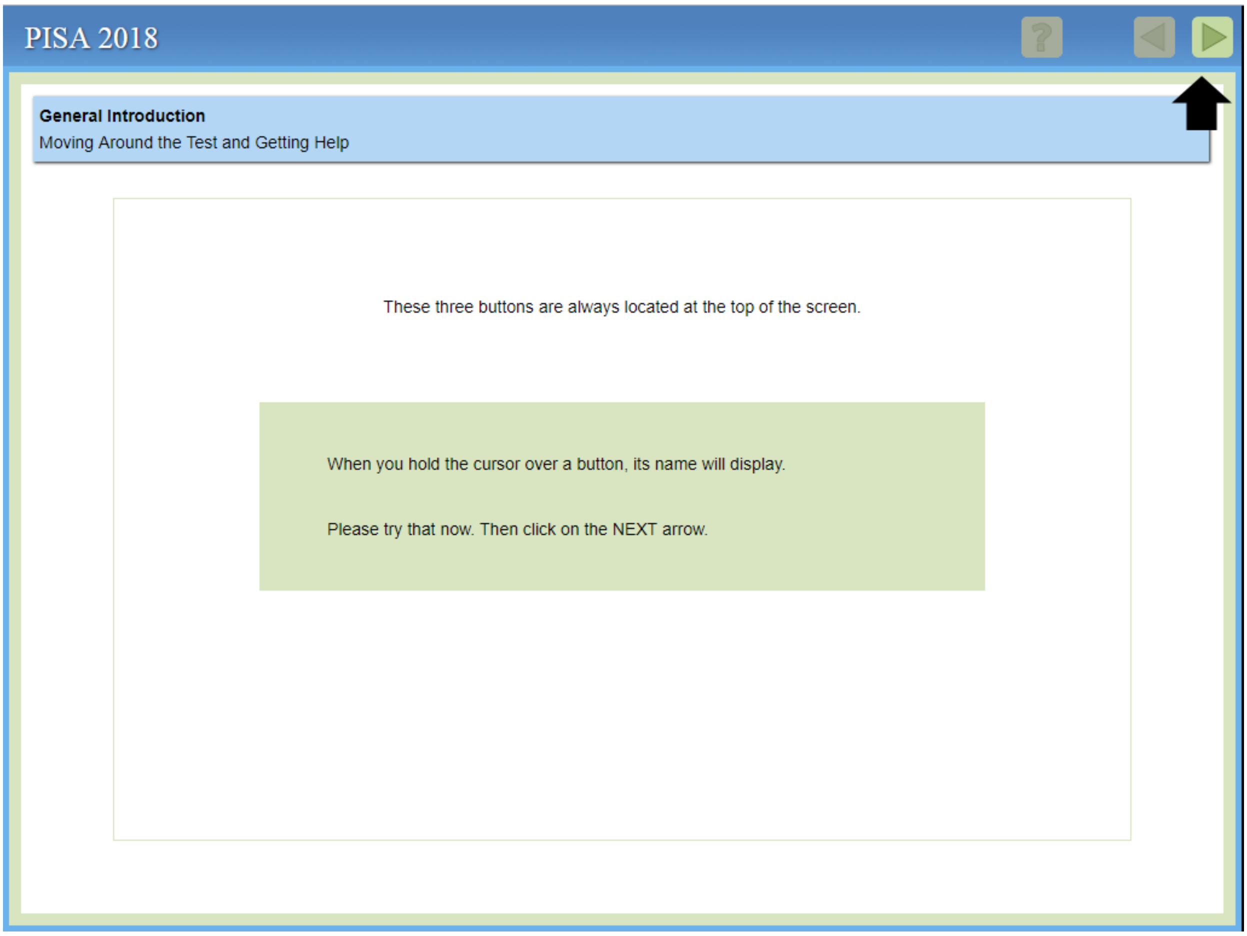


Figure B.2. Click on a choice (also referred to as a radio button or multiple choice)

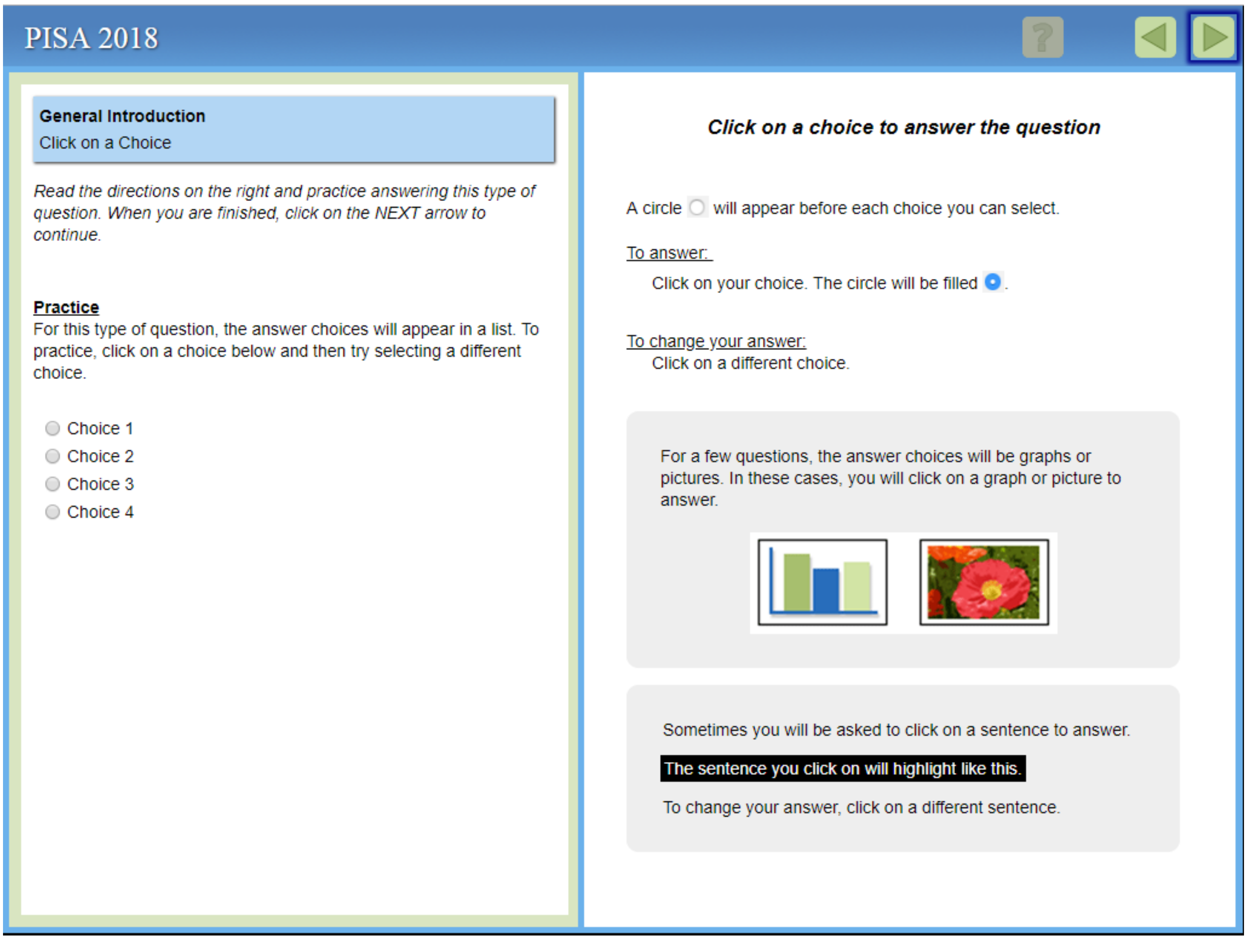


Figure B.3. Click on choices in a table (select answers among rows of choices in a table)

Type the subtitle here. If you do not need a subtitle, please delete this line.

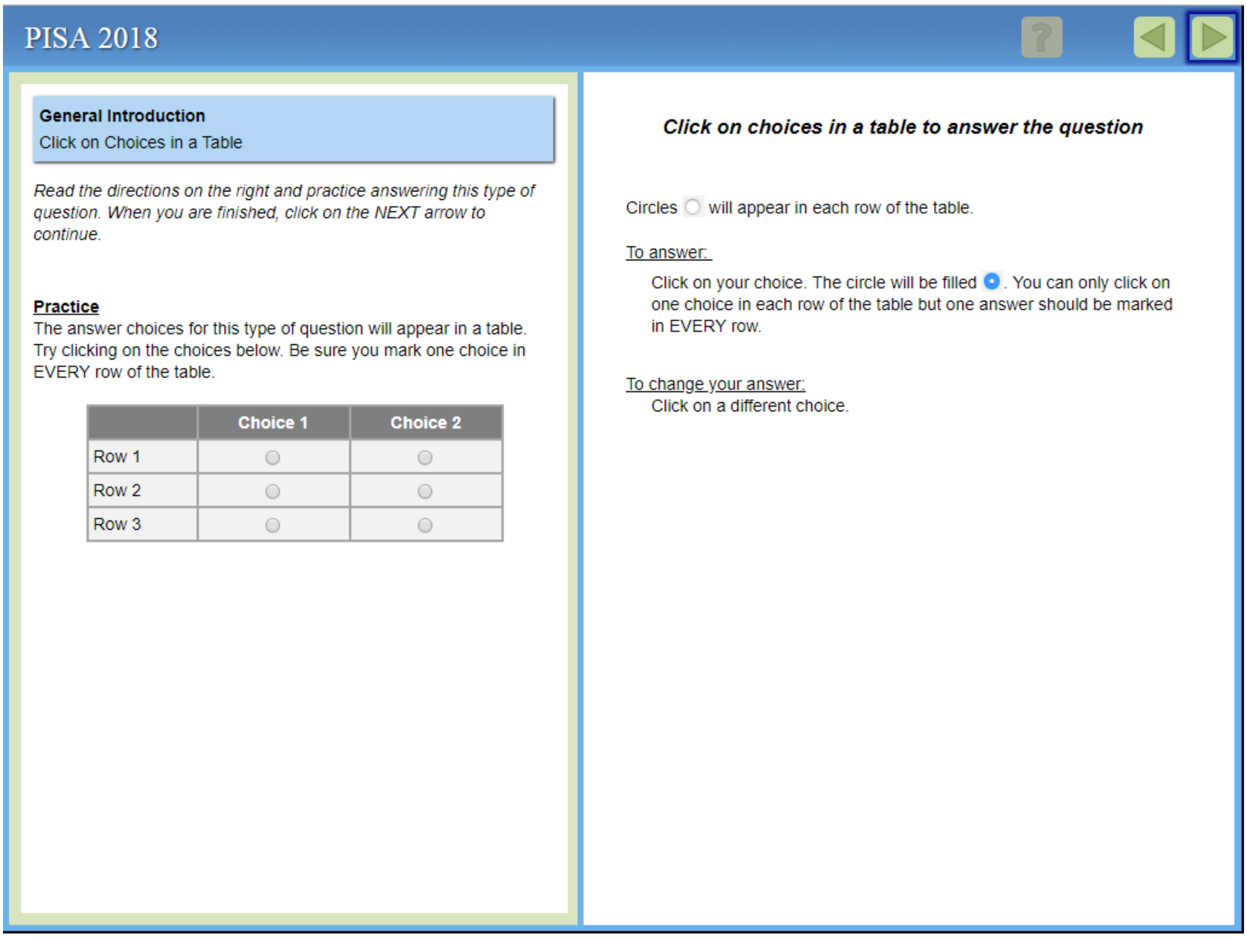


Figure B.4. 4. Click on one or more boxes (also referred to as a checkbox or multiple selection multiple choice)

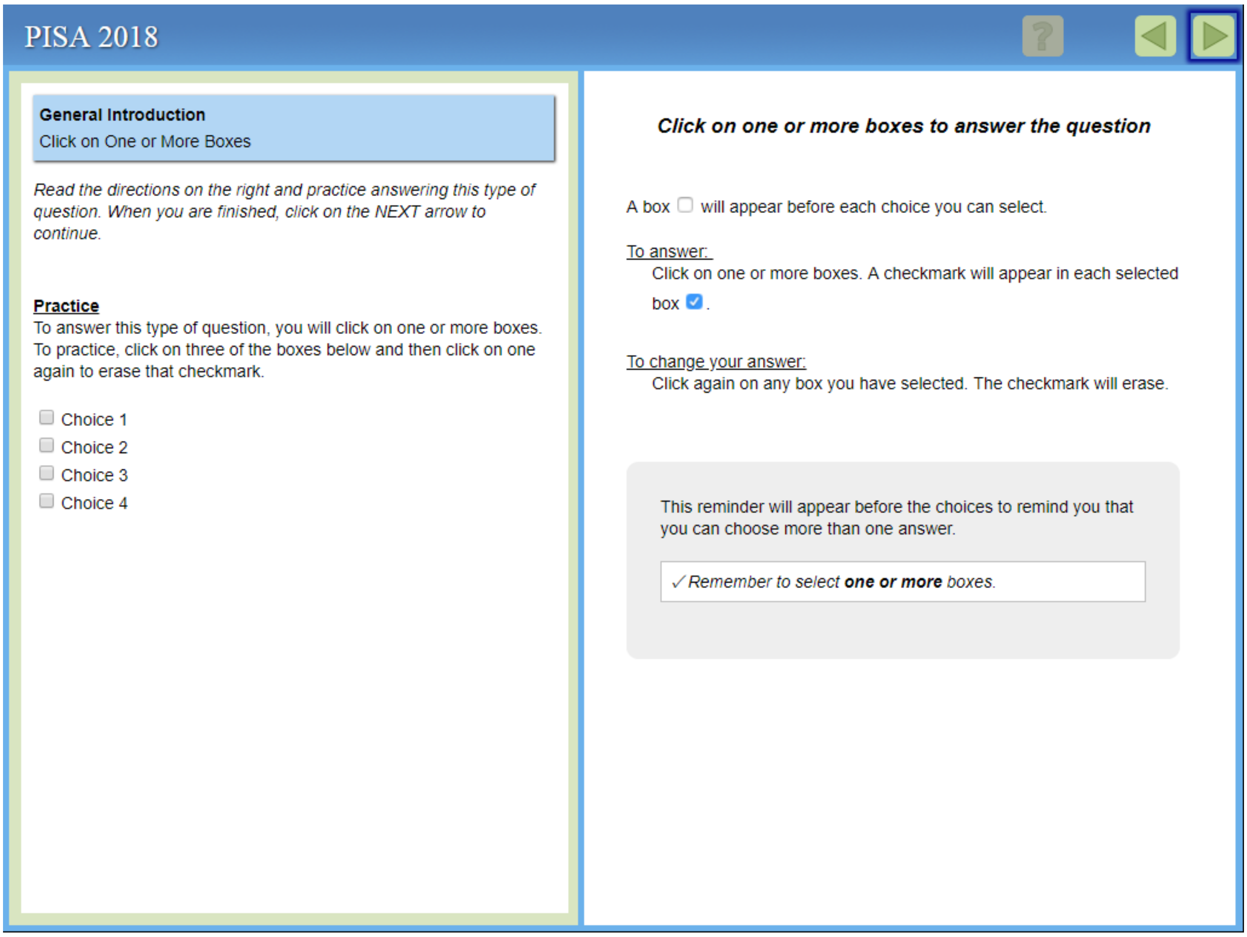


Figure B.5. Type text (also referred to as type your answer or open-ended text entry)

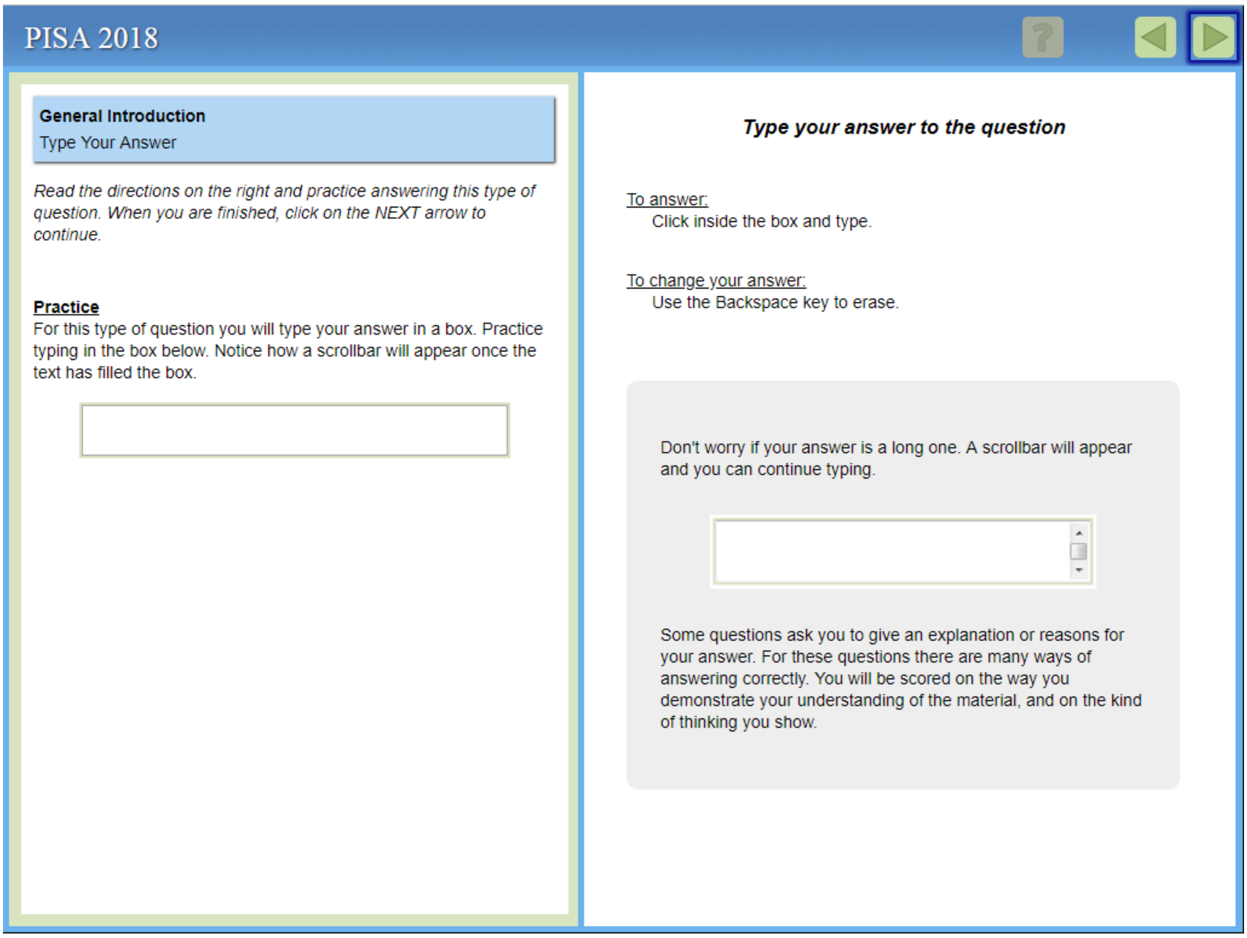


Figure B.6. Enter a number (open-ended numeric entry)

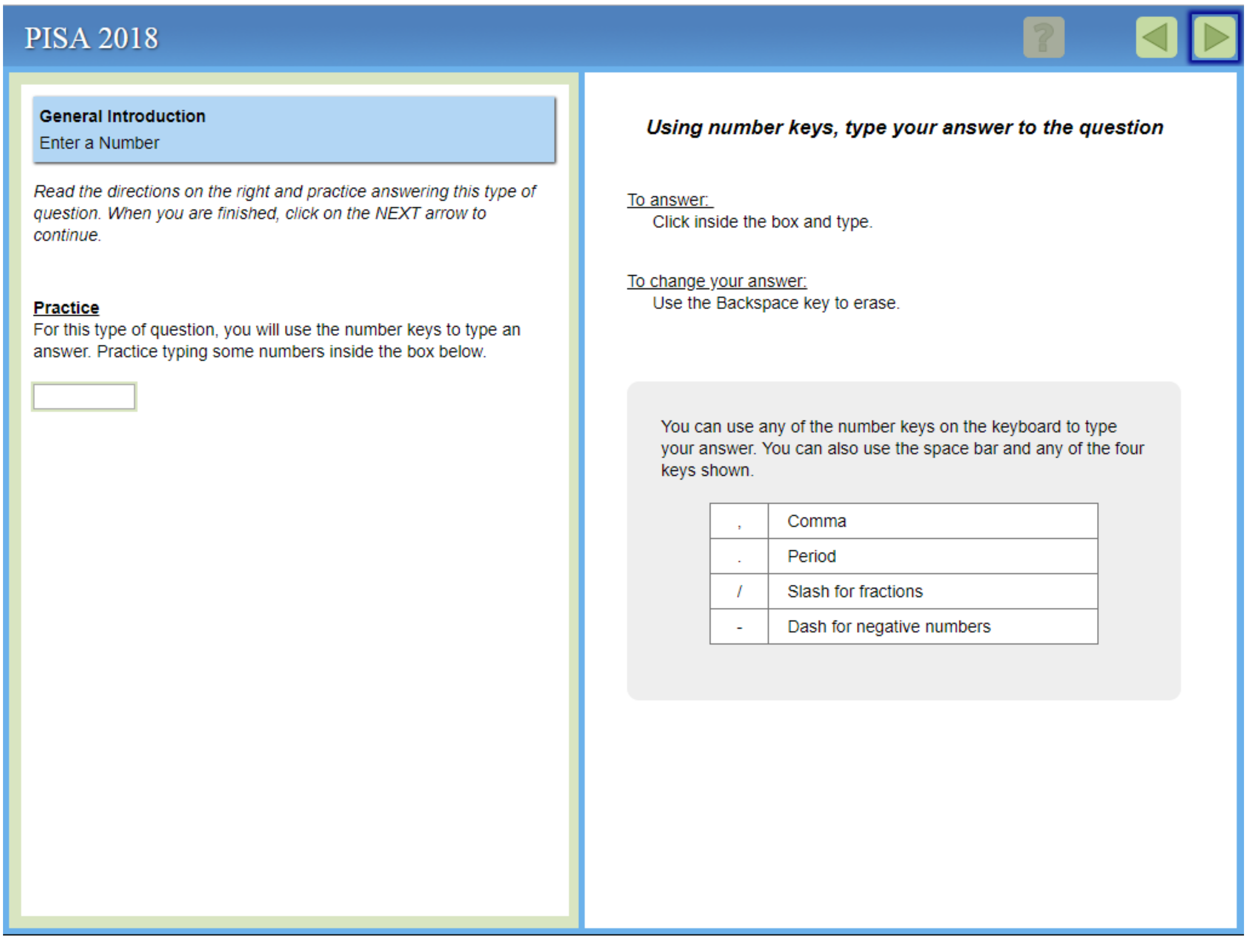


Figure B.7. Drop-down menus (select one answer per list of choices in a menu)

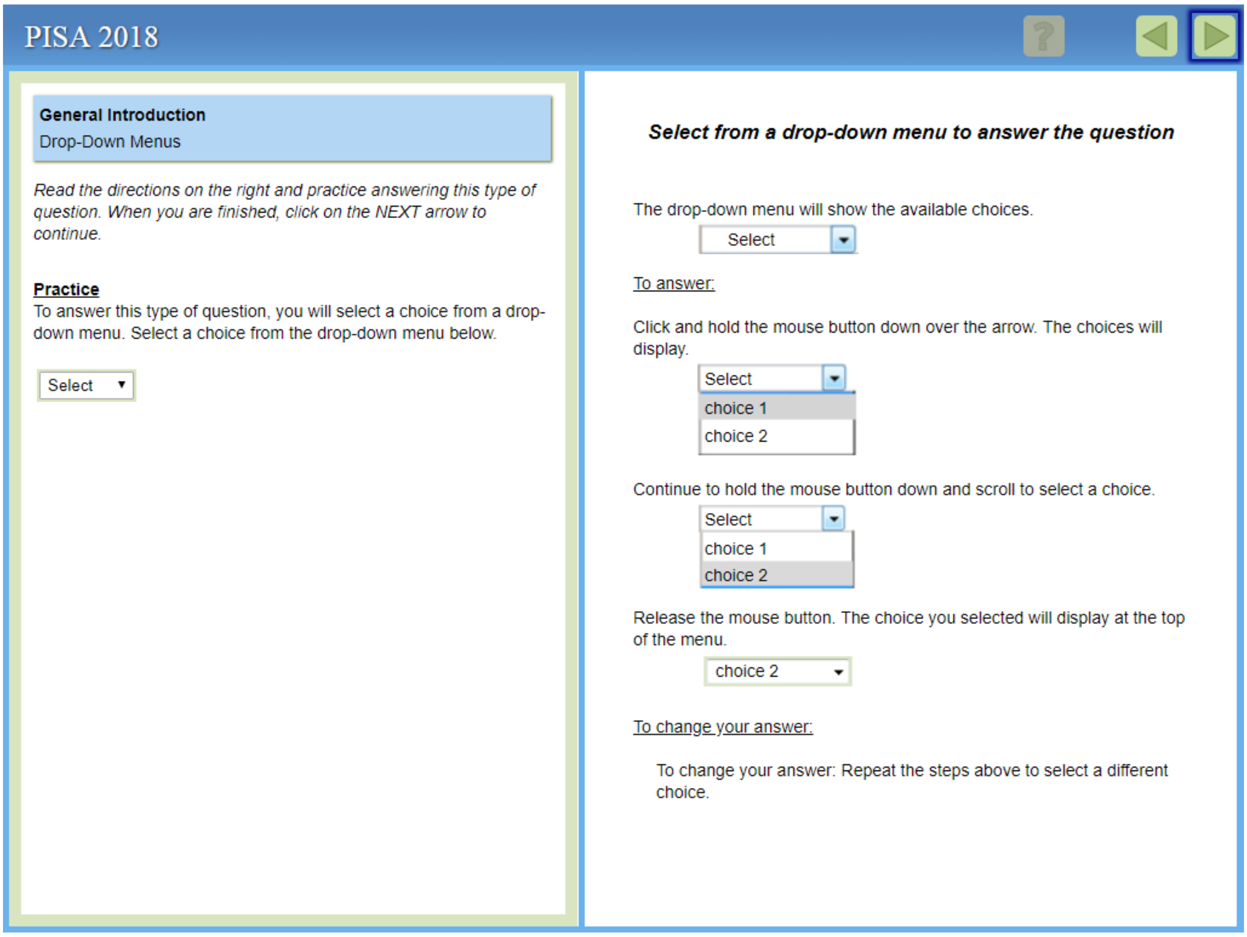


Figure B.8. Drag and drop (select an answer by moving an option from one column to a corresponding answer in another column)

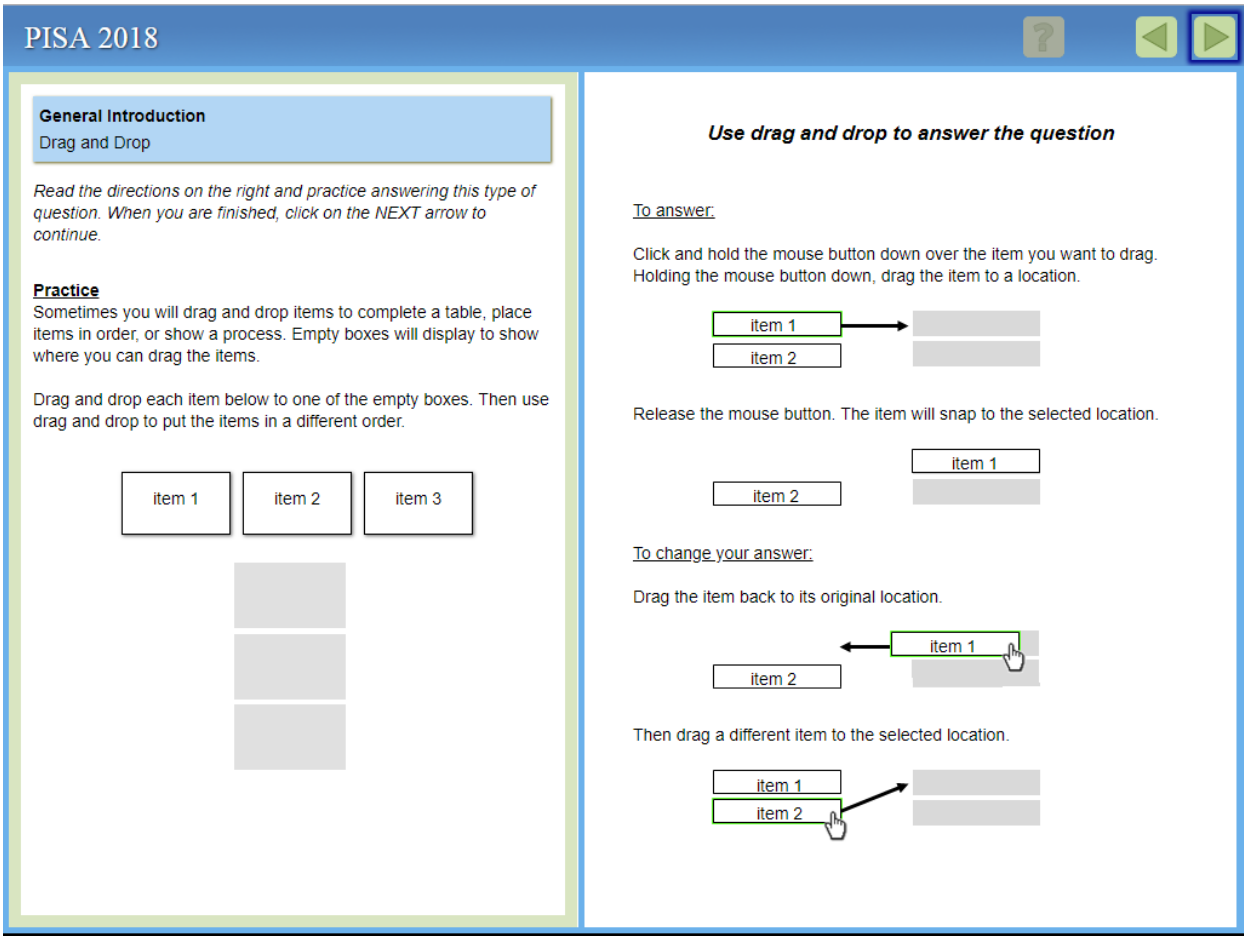


Figure B.9. Run a simulation (requires students to run a simulation to populate a data table)

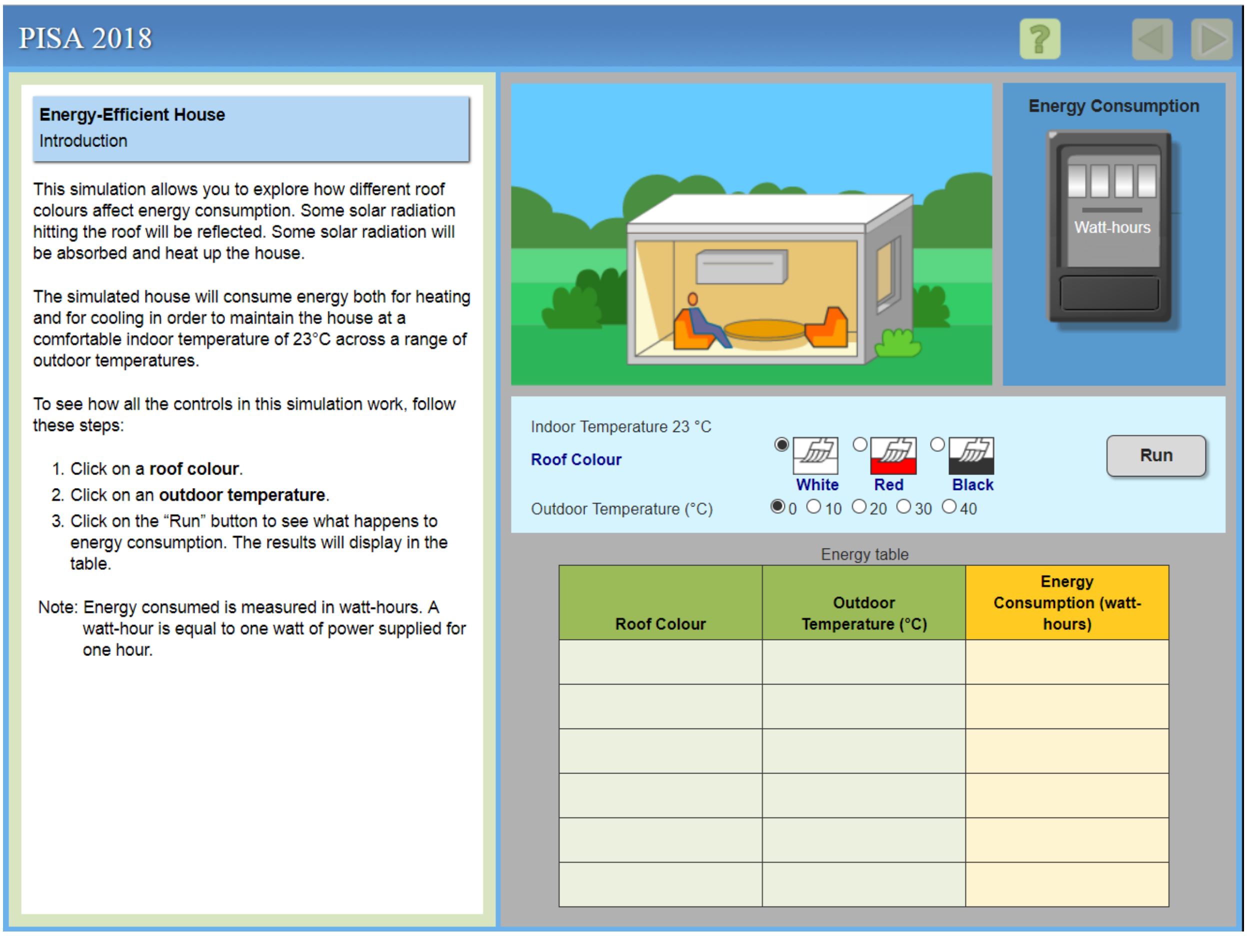
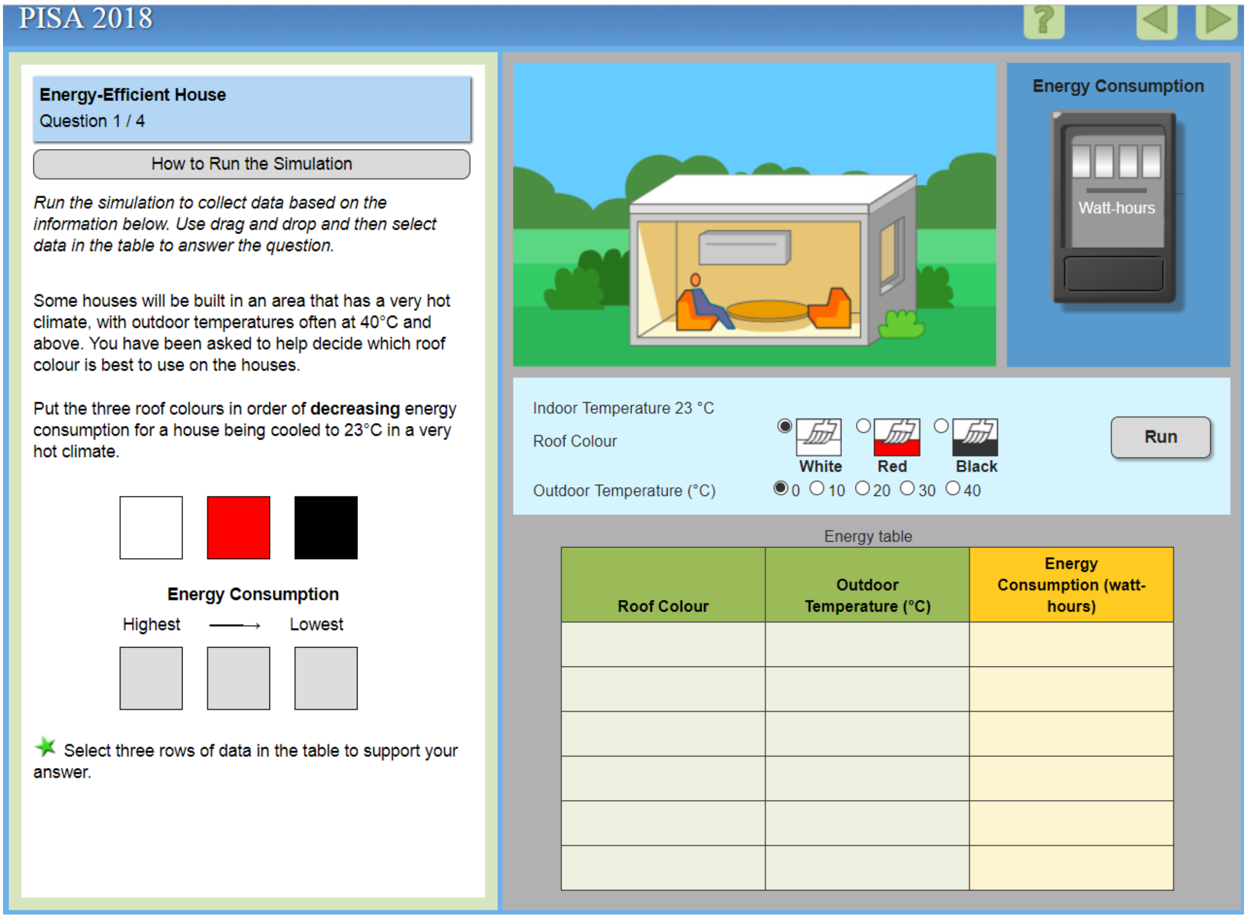


Figure B.10. 10.Combined task (drag and drop with a simulation)



Annex C. Changes to the PISA Orientation and Energy Efficient House Task to Adhere to WCAG 2.0

These are the changes that were made to the code of the PISA content so that it **was in compliance with WCAG 2.0**. The specific guideline that the change addresses is included with each description.

1) Navigation buttons updated to include ARIA role tags of "button" so that assistive technology can recognise custom navigation buttons as buttons. (4.1.2 Name, Role, Value)
2) The button to run the simulation was updated to include the "button" aria role. (4.1.2 Name, Role, Value)
3) Disabled buttons announced as being disabled. (4.1.2 Name, Role, Value)
4) Indoor temperature made selectable by mouse and keyboard. (2.1.1 Keyboard)
5) Controls to delete data from table were made accessible via keyboard. (1.3.1 Info and Relationships & 2.1.1 Keyboard)
6) Reading order for assistive technology updated to match the order of the visual presentation. (1.3.2 Meaningful Sequence)
7) Table interactivity (selecting, deselecting) was made accessible to assistive technology. (1. 3.1 Info and Relationships)
8) Table updated to include caption label. (1.3.1 Info and Relationships)
9) Alt text for image of energy efficiency house was revised to be more robust and meaningful. (1.1.1 Non-text Content)
10) How to run the simulation drop-down menu was updated to require users of screen readers to close the menu before being able to access the information behind the menu. (1. 3.1 Info and Relationships)
11) Drag-and-drop interaction was made keyboard accessible. (2.1.1 Keyboard)
12) Screen reader focus is limited to modal dialog boxes when they are present. (1. 3.1 Info and Relationships)
13) In the tutorial screens where tables present directions about how to use tools and other features found on the platform were given an ARIA role of "presentation" so that assistive technology knew not to present this as a data table. (1. 3.1 Info and Relationships)
14) Live radio buttons, drop-down boxes and checkboxes used as visual examples in the tutorial were updated to be static and to be ignored by assistive technology as they are intended to be supplemental and nonessential. (1.1.1 Non-text Content)

The following changes were made to the PISA content code that **go beyond the requirements of WCAG 2.0**. That is, ETS used our best practices in web content delivery for assessment that is not yet addressed in WCAG 2.0 so students can interact with test content and their assistive technology in a meaningful way.

1) Wattage meter presence suppressed for assistive technology because this information is presented in the data table.

2) Radio button interaction was updated to be in line with conventional radio button interaction. Navigation within a group of radio buttons can be done with the keyboard arrow keys, and the radio buttons are announced as part of a series.

3) Added a focus indicator that is easier to see than the default focus indicator found in the browser.

The following factors were identified as **features to change but were unable to be made** due to the time to revise the code and/or the project budget limitations.

1) Revise how questions are numbered so that a slash is not used. For example, use “1 of 4” instead of “1/4”.

2) Suppress iFrames when iFrames are empty to improve navigability. Provide the title to each iFrame. Also, add the frameset around the iFrame. More information can be found here: https://webaim.org/techniques/frames/

3) Change mouse-centric language to be applicable to users of keyboards.

4) Hide decorative images included in the tour of the platform from assistive technology as they do not add to the user’s experience.

5) Add directions regarding how to use drag-and-drop items with keyboard.

6) Add a legend for “Choose roof colour position” in drag-and-drop widget. Currently these legends are provided as a text that is not announced by screen readers.

7) Screen reader should announce, “Select three rows of data in the table support your answer.”

8) Provide the mechanism needed for screen readers to announce highest to lowest selected answer.

9) Screen readers should announce when a row has been added or removed from the data table.

10) Screen readers should announce instructions such as “Select two rows of data in the table to support your answer.”

11) Remove the document role from the “How to run simulations” button.

12) Wrap Question 4 /4 radio buttons in field-set grouping and provide the legend.

Finally there was one suggestion provided that may improve the overall usability for students using assistive technology. **This suggestion was not made** given the complexity of the work to make this modification.

1) Navigational buttons should be added to the bottom to move forward/backward for screen reader users. You may keep it visually hidden so that it is accessible only to screen reader users. This will give screen reader users easy access to navigate the application.