**Investigating the Usability of Technology-Enhanced Assessment Items During the ELPA21 Development Process**

Prepared by

Mikyung Kim Wolf, Danielle Guzman-Orth, & Jennifer Wain

Educational Testing Service

With additions regarding the use of results by ELPA21

ELPA21 Cognitive Lab Study Report

Final Deliverable

March 5, 2015

**Executive Summary**

This study was conducted as part of a large-scale contract supporting the English Language Proficiency Assessment for the 21st Century (ELPA21) Consortium, which is focused on the development of an English Language Proficiency (ELP) assessment system for K-12 English Language Learners (ELLs). At the time of this study, Educational Testing Service (ETS) was under a contract with the Council of Chief State School Officers (CCSSO) for the item development for the ELPA21 ELP assessment. CCSSO requested that ETS conduct a small-scale cognitive lab study to examine the quality of the items during this development stage, particularly for their new technology-enhanced (TE) features. This was part of ELPA21’s efforts to ensure that the new, innovative item types function appropriately for students to demonstrate their English language proficiency. The results of the study were expected to provide useful information for future development of the test platform and administration guidelines before a large-scale field-test or operational testing for the ELPA21 assessments.

Specifically, the purpose of the study was to examine students' interaction with new TE item types and use of technology features during the computer-based testing[[1]](#footnote-1). The areas of investigation focused on the clarity of item directions, the usability of technology features (e.g., drag & drop, zone selection, recording functions, mouse, keyboarding, etc.), and the usability of accessibility tools (e.g., a line reader, a magnifier, a text reader, etc.). With the assistance of the CCSSO and the Washington State Office of Superintendent of Public Instruction (OSPI), a total of 91 students in Grades K-3 and Grades 6-8 were recruited from two elementary schools and one middle school in two urban districts in Seattle. The participants included current ELL students at different ELP levels, fluent English-speaking students (non-ELLs and reclassified ELLs), and ELL students with disabilities with an Individualized Education Program (IEP)/504 Plan. The primary language backgrounds of the ELL students in the sample were Spanish, Russian, and Vietnamese.

A trained ETS team consisting of researchers and assessment specialists conducted a series of one-on-one cognitive labs over three days in October 2014. The interviewers’ detailed observation notes, ratings, students’ verbal interview data, and students’ test performance data were qualitatively analyzed to summarize the patterns of student performance on the TE items and computer-based test platform.

Overall, the study yielded promising results for the quality of the item types and use of technology features. Below is a summary of a few major findings:

* Students were able to quickly learn technology features and complete the TE item types on a computer when clear directions were provided.
* Students were highly engaged and interested in the items. This observation was particularly evident in younger grades (i.e., Grades K-2).
* Students in Grades K and 1 needed individual assistance to complete the TE items during the computer-based testing. Students with a beginning level of ELP needed extra support in understanding the directions across all the grade levels.
* ELL students with disabilities (ELSWDs) were mostly able to interact with the TE features after receiving additional assistance. Some ELSWDs (i.e., students with speech disabilities, visual impairment) needed more individual assistance than others during the computer-based testing.

Despite the limitations of the study, namely a small sample size, the study findings offered valuable information to enhance the quality of the ELPA21 assessments. Based on these findings, the following recommendations and suggestions are provided for future development or administration of the operational ELPA21 assessments, followed by how these recommendations and suggestions were addressed in the field test and or will be addressed post field test[[2]](#footnote-2).

* A tutorial and sample practice items should be provided to the students prior to testing.
  + ELPA21 provided an interactive item demonstration site for the field test, accompanied by detailed lesson plans, for each of following the grade bands: K, 1, 2-3, 4-5, and combined 6-8 and 9-12. The demonstration items covered each of the item interactions in the grade-level texts. Teachers were encouraged to both demonstrate the items to their students and provide their students with opportunities to practice with the items so they would be familiar with the testing environment.
* Clear and explicit directions must accompany each item type.
  + Each item type on the field test is accompanied by specific directions. Student responses and reports from field test administrators will be reviewed to determine the effectiveness of the instructions.
* The construct measured in each item type should be clearly defined in order to determine any necessary accommodations or specific accessibility features for a computer-based test.
  + The *Accommodations and Accessibility Framework* that guides ELPA21 takes the constructs being assessed into account, and items and tasks were developed with available accommodations and accessibility tools in mind. Student responses and reports from field test administrators will be reviewed to determine whether the administration procedures are appropriate, particularly attending to the Writing issues noted in this report.
* Accessibility tools designed to support all students’ access to the test content should be designed to work with innovative TE item types and the computer platform; the tools must be demonstrated to all students prior to testing.
  + The interactive item demonstration described above includes the accessibility tools that were used on the field test.
* Individual assistance should be provided during the test administration for younger students (e.g., students in Grades K-2).
  + The recommended administrator-to-student ratio for Grades K and 1 for the computer-based portion of the assessment is one administrator for one to five students, and for the paper-based portion it is one administrator for one to three students; for Grade 2 it is one administrator for one to eight students. Reports from field test administrators will be reviewed to determine whether this ratio is sufficient to allow individual assistance for these students.
* ELSWDs should be provided with individual assistance during test administration, as well as sufficient opportunities to practice using the TE features and platform prior to testing.
  + The interactive item demonstration site provided practice with designated accessibility features and embedded accommodations for students who would use them during the field test. In addition, the *Test Administration Manual* provides information about what type of help can be provided to students during test administration, such as assistance with scrolling, clicking, and starting/stopping the recordings.
* Item formatting should be clear and intuitive and should not require unnecessary use of technology navigation features.
  + Improvements were made to the field test platform based on the cognitive lab results, and more are planned for the operational platform. In addition to the cognitive lab results, reports from field test administrators will be reviewed to determine what changes are needed.
* Technical requirements should include detailed specifications about the equipment and must be tried out prior to testing.
  + ELPA21 developed a *Setup and Installation Guide* that includes detailed specifications for hardware and bandwidth tryout guidance, and tips for making sure the testing system was implemented properly.

The results of this study signify the importance of further empirical research. One area of research should investigate whether future modifications made to address some of the issues reported in this study improve the testing experience for students. Sustained research will be needed to monitor the appropriate use of the technology features and provide useful information to enhance the validity and technical quality for the ELPA21 assessments.

**Background of the Study**

The English Language Proficiency Assessment for the 21st Century (ELPA21) Consortium aims to develop an assessment system that provides high quality measures of English language learner (ELL) students’ progress in the development of their academic English language proficiency (ELP). Based on the ELP Standards adopted by ELPA21, the screener and summative ELP assessments in this system are intended to reflect important language skills that ELL students need to acquire in order to meaningfully participate in English-medium instruction settings and be prepared for college and career success.

Coupled with the new ELP standards, which correspond to the language practices embodied in the college and career readiness standards (e.g., Common Core State Standards, Next Generation Science Standards), ELPA21’s assessments feature the use of technology to develop and deliver innovative and engaging item types on a computer, often in a manner intended to more closely reflect the target constructs than can be done using paper and pencil. Devising appropriate and engaging item types that allow students to demonstrate their English language proficiency is a critical factor to increase the validity of score interpretations for the assessments. ELPA21 also intends to integrate various accessibility tools (e.g., highlighter, line reader, magnifier, text reader, etc.) into the assessments with the use of technology, thereby making the accessibility features available to all students in need. Considering that students are becoming increasingly exposed to the use of digital tools in their learning environments, technology-enhanced (TE) assessments seem necessary to keep pace with current educational practices.

In contrast, it is equally important to ensure that the use of technology does not inadvertently cause any systematic construct-irrelevant variance in assessing ELL students’ ELP. For instance, the directions for TE items (e.g., clicking a zone, drag & drop, drop-down menu) should be clear for students who may be accustomed to traditional paper-based assessments. As the Standards for Educational and Psychological Testing (AERA, APA, NCME, 2014) emphasize, fairness and validity need to be carefully considered throughout all stages of test development and use.

To date, only a few states have implemented a computer-based ELP assessment with their K-12 ELL students. The majority of the ELPA21 consortium states will newly implement the computer-based ELPA21 assessments for their ELL students. Prior to conducting a large-scale field-test, ELPA21 decided to conduct a small-scale trial study to examine students’ interaction with the TE features embedded in various item types. A close examination of students’ use of technology features was expected to provide valuable information to prepare for the administration guidelines as well as the design of the test platform for the large-scale field test and future operational assessments. Furthermore, the study findings would offer an important piece of empirical validity evidence to support the appropriate use of TE items and undergird the rigorous process undertaken during the ELPA21 assessment development.

The present document reports on the small-scale trial study that a team of researchers and assessment specialists from ETS conducted to address ELPA21’s needs. In particular, the study focused on the usability of the new item types and students’ behaviors during the computer-based testing. In collaboration with the ELPA21 Task Management Teams, the ETS research team designed a cognitive lab study, which was implemented with the assistance of one of the consortium states. Questar Assessment, Inc. (QAI), as a partner to aid in a large-scale field testing for ELPA21, was also involved in this trial study by providing the test delivery platform. Given the purpose of the study, a cognitive lab method was utilized wherein a trained researcher closely observes and interviews individual test takers on a one-on-one basis during testing. This technique has been prevalently used for investigating individual test takers’ behaviors in a systematic way (e.g., Martiniello, 2008; Winter, Kopriva, Chen, & Emick, 2006; Wolf, Kim, & Kao, 2012). It should be noted that the study was exploratory in nature and conducted with a small sample size to identify any potential issues for further investigation, modification or recommendations for the use of TE features and assessments.

**Research Questions**

The study primarily focused on three areas that pertained to the technology features: (1) clarity of directions, (2) students’ interaction with TE items, and (3) accessibility features embedded in the preliminary test platform. Note that the built-in accessibility tools were available to all students. Within each area, the study addressed the following research questions:

1. Clarity of directions for students

* Are the directions clear enough that students can tackle the items independently?
* What difficulties do the students at different levels of English proficiency have in understanding the directions, if any?
* What difficulties do the students with disabilities have in understanding the directions, if any?

1. TE item interaction (including computer use)

* To what extent can the students complete new item types (particularly TE items) independently?
* What difficulties do the students have in completing TE items?

1. Accessibility for all students and ELLs with disabilities

* To what extent and how do the students use the accessibility tools?
* What types of difficulties do students experience in accessing and interacting with the accessibility features (i.e., universal and designated features)?
* To what extent do the accessibility tools meet the needs of students with disabilities while they are interacting with the assessment items?

**Method**

**Participants**

With the assistance of the Washington State Office of Superintendent of Public Instruction (OSPI), a total of 91 students in Grades K-3 and Grades 6-8 were recruited from two elementary schools and one middle school in two urban districts in Seattle. Due to time and resource constraints, not all of the grade levels were included in this study. Yet, efforts were made to maximize the usability of the study findings by selecting younger grades rather than older grades for similar item types. For instance, since the TE item types for Grades 4-5 were similar to those for Grades 2-3, we focused on examining younger grade students’ interactions with the TE item types with an assumption that younger grade students may have more difficulties with TE features. Similarly, Grades 9-12 were not included as their item types were similar to those for Grades 6-8.

To investigate the clarity of the directions and the use of technology features in specific item types, we attempted to include ELL students at different ELP levels as well as fluent English speakers (i.e., native English speakers and reclassified ELLs). For the purpose of this study, we categorized the latter group as non-ELL in comparison to our focal group of ELL students. The inclusion of non-ELL students was important in determining whether the sources of difficulties students might have are related to ELL-specific issues. For instance, if non-ELL students who are fluent in English display some difficulty in understanding the directions, this would be an indication of the need to further clarify the directions. As a small-scale study, we limited our ELL participants’ home language background to Spanish, Vietnamese, and Russian. These were the top three languages spoken by the ELL population in Washington and were commonly included in other ELPA21 states. We also made efforts to include ELL students with disabilities in order to examine the usability of the accessibility tools and appropriateness of the new item types for these students. Table 1 summarizes the number of participants for this study.

Table 1. The Number of Participants by Grade, ELP Levels, and Disability Status

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ELP Levels and Disability Status | Grade or Grade Bands | | | |  |
| Grade K | Grade 1 | Grades 2-3 | Grades 6-8 | Total |
| Beginning ELL | 5 | 1 | 2 | 2 | 10 |
| Intermediate ELL | 9 | 7 | 8 | 15 | 39 |
| Advanced ELL | 0 | 3 | 7 | 0 | 10 |
| Reclassified ELL | 0 | 2 | 4 | 3 | 9 |
| Native speakers of English | 6 | 2 | 1 | 0 | 9 |
| ELL with disabilities | 4 | 1 | 3 | 6 | 14 |
| Total | 24 | 16 | 25 | 26 | 91 |

In our sample, ELL students with an Individualized Education Program (IEP)/504 Plan represented a range of high incidence disabilities (speech language impairment, specific learning disability) and low incidence disabilities (other health impairment, visual impairment). Other disability characteristics included behavioral, social, motor difficulties, and other health impairments. Several students had multiple disabilities (e.g., speech and social; speech, behavior, reading, writing, math; speech, motor, reading, writing, math). Table 2 summarizes the characteristics of the ELL students with disabilities (ELSWDs) in our study.

Table 2. The Characteristics of the ELL Students with Disabilities in the Study Sample

|  |  |  |  |
| --- | --- | --- | --- |
| Student # | Grade or  Grade Bands | Disability Categories | ELP Levels |
| S1 | Grade K | Speech | Beginning |
| S2 | Grade K | Speech | Beginning |
| S3 | Grade K | Speech, social | Beginning |
| S4 | Grade K | Speech, reading, writing, math, behavior | Beginning |
| S5 | Grade 1 | Speech, social, reading, writing, math | Intermediate |
| S6 | Grades 2-3 | Reading, writing, math, social | Advanced |
| S7 | Grades 2-3 | Reading, writing, math, social | Advanced |
| S8 | Grades 2-3 | Speech, OT (motor), reading, writing, math | Advanced |
| S9 | Grades 6-8 | Visual impairment (late onset) | Advanced |
| S10 | Grades 6-8 | Other health impairment | Intermediate |
| S11 | Grades 6-8 | Specific learning disability | Beginning |
| S12 | Grades 6-8 | Specific learning disability | Intermediate |
| S13 | Grades 6-8 | Specific learning disability | Advanced |
| S14 | Grades 6-8 | Specific learning disability | Advanced |

Note: Disability was defined as an existing IEP/504 plan; however, the disability label displayed in the disability category column is the language of the teacher or site coordinator and does not reflect the formal federal disability categories established in IDEA (2004).

**Study Instruments**

The ETS research team developed a cognitive lab protocol that the interviewers could use to collect their observations and verbal data from each student. The protocol contained a set of interview questions to examine students’ understanding of the directions and any difficulties in using technology features in the item types. The protocol also included a set of rating scales to evaluate students’ understanding of the directions, use of TE features, use of the accessibility tools, engagement in the items, and independence in completing the items (see Appendix A for the interview questions and rating scales). The interviewers were also asked to write down comments to justify their ratings based on their observations and interviews.

In addition, a student background questionnaire was developed to collect the participating students’ background characteristics such as home language, ELL status, ELP levels, reading proficiency levels according to other testing measures, and the types of disabilities, if applicable. The questionnaire was completed by the teachers from each school. All of the participants were assigned unique IDs created specifically for this study to protect anonymity.

**Item Types in the Study**

Due to the limited time (i.e., one class period of time for each cognitive lab), item types were purposefully selected from each listening, speaking, reading, and writing domain. The primary criterion for selection was to include the new item types with TE features (e.g., drag & drop, drop-down menu, zone selection, recording functions, etc.). Additionally, these item types were selected to examine the students’ use of the computer, including a mouse and keyboard. Tables 3 and 4 present a summary of the selected item types for elementary and middle school grades, respectively.

Table 3. Item Types Selected for the Study, Grades K-3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Domain | Item Type | Grade K | Grade 1 | Grades 2-3 |
| Listening |  |  |  |  |
|  | Listen and Match |  |  |  |
|  | Follow Instructions |  |  |  |
|  | Short Conversation |  |  |  |
|  | Long Conversation |  |  |  |
|  | Teacher Presentation |  |  |  |
|  | Read aloud Story |  |  |  |
|  |  |  |  |  |
| Speaking |  |  |  |  |
|  | Classroom Tableau |  |  |  |
|  | Show and Share Questions |  |  |  |
|  | Show and Share Presentation |  |  |  |
|  | Picture Description |  |  |  |
|  | Conversation |  |  |  |
|  | Observe and Report |  |  |  |
|  | Opinion |  |  |  |
|  |  |  |  |  |
| Reading |  |  |  |  |
|  | Word Wall |  |  |  |
|  | Read along Story |  |  |  |
|  | Read and Match |  |  |  |
|  | Short Correspondence |  |  |  |
|  | Read along Sentence |  |  |  |
|  | Read for Details |  |  |  |
|  | Informational Set |  |  |  |
|  | Literary Set |  |  |  |
|  | Procedural Text |  |  |  |
|  |  |  |  |  |
| Writing |  |  |  |  |
|  | Word Builder |  |  |  |
|  | Sentence Builder |  |  |  |
|  | Picture Caption |  |  |  |
|  | Write an Opinion |  |  |  |

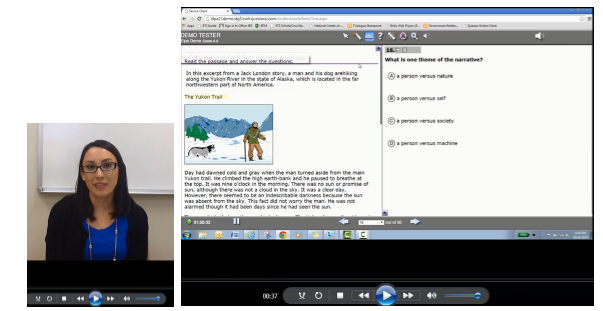
Table 4. Item Types Selected for the Study, Grades 6-8

|  |  |  |
| --- | --- | --- |
| Domain | Item Type |  |
| Listening |  |  |
|  | Listen for Information |  |
|  | Academic Lecture or Discussion |  |
|  | Interactive Student Presentation  Academic Debate |  |
|  |  |  |
| Speaking |  |  |
|  | Compare Pictures |  |
|  | Analyze a Visual and a Claim |  |
|  | Language Arts Presentation |  |
|  | Analyze a Visual and a Claim |  |
| Reading |  |  |
|  | Short Informational Set |  |
|  | Short Literature Set |  |
|  | Extended Informational Set |  |
|  | Extended Literature Set |  |
| Writing |  |  |
|  | Storyboard |  |
|  | Discrete Editing Tasks |  |
|  | Writing Questions Tasks  Responding to a Peer E-mail  Construct a claim |  |

**The Test Platform for the Study and a Tutorial**

For this study, a cognitive lab test platform was created by QAI to deliver the selected items as a test form for each grade or grade band on a computer. By participating in the study, QAI aimed to learn about any potential issues with the platform prior to a large-scale field test for ELPA21. Thus, the design of the test platform was a preliminary one, mimicking the future field-testing platform. It was expected that the study results would provide useful information to finalize the operational platform design for delivering the items on a computer. The test platform for this study also contained a set of built-in accessibility tools (e.g., digital notepad, highlighter, line reader, magnifier, text reader).

At the time of this study, a tutorial or demo of sample items was not yet available. Since it was important to explain the test layout, major icons, and functions (e.g., recording buttons, audio buttons, and accessibility tools) to the test takers, the ETS team created a short tutorial video for the study. The Grade K and 1 version was 3:47 minutes long, and the Grade 2-3 and 6-8 version was 4:20 minutes long. The tutorial was necessary not only for demonstrating the item layout and features but also for standardizing the cognitive lab conditions. Figure 1 displays a screenshot of the tutorial video.



*Figure 1.* Two screenshots of a tutorial video created for the study.

Since the test platform for this study was a preliminary one, some general instructions before each domain section (i.e., listening, reading, speaking, and writing) were also unavailable.

**Procedure**

A total of nine interviewers received two 2-hour training sessions to conduct the cognitive labs according to the protocol. The majority of the interviewers also had prior experience in conducting cognitive labs with K-12 students. Some interviewers were fluent speakers of Spanish, Russian, or Vietnamese. in order to conduct interviews in those languages when needed. In total, 17 out of 91 interviews were either conducted entirely in the students’ home language (14 students) or assistance was provided in the home language when necessary (3 students).

The cognitive labs were conducted over three days from October 21, 2014 to October 23, 2014 at the school sites. Note that the study was conducted at an earlier time within the school year than the operational summative test will be administered. Thus, kindergarteners were still relatively new to school, and most of the Grade 1 students in the sample could not read a sentence-level text yet. This factor needs to be taken into consideration in interpreting the study results.

Each one-on-one cognitive lab was conducted with a laptop (14 inch screen) and a standard-sized mouse (except one interviewer who used a smaller-sized mouse). Each student used a headset with a microphone attached. Before beginning the test, each student watched the short video tutorial showing how to use the accessibility tools as well as the audio buttons, recording buttons, arrows, and scroll bar. Several questions pertaining to the students’ use of and familiarity with computers were also asked prior to beginning the items. The interviewer asked the questions from the cognitive lab protocol after each item type or each domain, depending on the student’s pace of answering the questions. During the test, the interviewer also explained the

TE features (e.g., the use of the audio buttons, recording buttons, arrow buttons, scroll bar, mouse) when the student had difficulty understanding the features needed for answering the question.

All of the cognitive lab interviews were audio-recorded. The interviewers completed their ratings in the areas of the student’s (1) understanding of directions, (2) degree of difficulty in the use of TE features, (3) independence in solving the questions, (4) engagement, and (5) use of accessibility tools. The interviewers also took detailed notes following the protocol. After completing the cognitive labs each day, the interviewers had a debriefing meeting to share their notes and immediate patterns they observed. The interviewers’ notes and ratings were used as the data set, along with the students’ responses.

**Analysis**

Descriptive statistics for the student background data and the interviewers’ ratings were computed. For the areas that received low ratings, the student background, student responses, and interview notes were reviewed to find out the source of the low ratings. The interviewers’ detailed notes were qualitatively analyzed. Four researchers did multiple readings of the notes and categorized them according to the areas of interest in this study. Commonly identified patterns of student performance and areas of difficulty were then summarized. To accurately distinguish the source of the item difficulties, which could have possibly stemmed from students’ disability types, the analysis and qualitative summaries were conducted separately for students with disabilities.

**Results**

We first present the results for ELL and non-ELL students regarding the three major research questions: clarity of the directions, the students’ interaction with TE features, and the use of the accessibility tools. Then, we present the findings from ELL students with disabilities, focusing on their interaction with the items and TE features, including the accessibility tools. The results are summarized by grade level. Some detailed findings about students’ interaction with specific item types are included in Appendix B.

**Grade K**

Table 5 presents the interviewers’ overall ratings in the five major categories on a 3-point scale across the item types. The sample included a total of 20 students without disabilities (14 ELLs and 6 non-ELLs). As seen in Table 5, only 10% (2 students) of the sample were rated to have completed the items independently. A closer look at the sample indicated that the two students were ELL students with intermediate ELP, suggesting that even non-ELL students had some difficulties in completing the items on their own. The interviewers’ notes suggested that the students at this grade level had a strong tendency to check with the interviewer regarding their understanding of the directions or ask for help with the use of TE features. In addition, there was a frequent need for the interviewers to initially prompt the students to speak for the speaking items. Yet, many students expressed their enjoyment and interest in working on the items, particularly for the “Picture Description” item type where students were expected to find and describe strange things in a picture.

Table 5. The Frequencies of the Overall Ratings across Item Types: Grade K

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Freq. (%)* | | |
| Categories | 0  no | 1  partially | 2  entirely |
| Understanding the directions | 4 (20) | 9 (45) | 7 (35) |
| Difficulties in the use of TE features | 8 (40) | 11 (55) | 1 (5) |
| Completing the item independently | 5 (25) | 13 (65) | 2 (10) |
| Engagement in the items | 0 (0) | 4 (20) | 16 (80) |
| Using the accessibility tools | 19 (95) | 1 (5) | 0 (0) |

***Clarity of the directions.*** The interviewers’ ratings in Table 5 showed that many students had some difficulty in understanding the directions across the items on average (65% being at the scale point of 0 or 1). A closer look at the interviewers’ notes reveal that this rating was largely attributed to the following major factors: (1) a lack of explicit directions about the test platform functions (e.g., arrow buttons used to progress to the next item, audio buttons, recording buttons), (2) Grade K students’ limited experience with testing overall, and particularly with a computer-based speaking test, and (3) some ELL students’ low English proficiency. Almost all of the students did not know how to progress to the next item after the first item. It is interesting to note that the arrow buttons that might be self-explanatory to experienced test takers or adult learners were not obvious to the kindergarteners who had limited experience with computer programs.

In addition, many students had some difficulty in understanding when to click or drag & drop. This was particularly the case for the “Word Wall” item type of the reading domain. The interviewers’ notes indicate that 45% (9 out of 20 students) of the sample were not clear as to whether to click a zone or drag & drop the choices. In another example from the speaking domain, many students tried to click on a picture in the “Show and Share Presentation” item. The students wanted to indicate their selection first before talking about the picture. An examination of the interviewers’ ratings for each item type also indicated that all of the students did not understand how to use the recording buttons for the first speaking item. Despite the video tutorial, the students at this grade level needed more explicit directions and opportunities to practice for the new item type. These results highlight the importance of explicit directions for the students in Kindergarten who have little experience with testing. Finally, some students at the beginning level had difficulties understanding the directions due to their lack of English language proficiency. The interviewer’s demonstration on how to use the TE features (e.g., drag & drop, recording buttons) was helpful for these students to understand the directions.

***Use of TE features in new item types.*** The interviewers provided their ratings for the students’ use of TE features, particularly for drag & drop, zone selection, and recording buttons at this grade level. In the sample, 60% of the students were rated to have some difficulties with these TE features at least once across the items. The interviewers’ notes indicated that the mouse was the major obstacle for the students in performing a drag & drop. That is, the standard-sized mouse was too big for the kindergarteners in this sample. Often times, students had difficulties with simultaneously holding and pressing the mouse for the drag & drop, particularly in the case of a relatively long distance between the object and the dropping zone. According to the interviewers’ notes, 45% of the sample were found to have this drag & drop problem as a result of difficulty in using the mouse. The interview data also revealed that 40% of the sample had never used a computer before. This fact might have contributed to the students’ interaction with the TE features.

A common pattern also emerged from the speaking item types, which required the students to say their response aloud after pressing the recording button. Almost all students needed to learn how to use the three recording buttons (i.e., record, stop, and playback) in the speaking items. As described above, the problem with the recording buttons seemed mainly due to a lack of practice or explicit directions on the use of the three buttons. It should also be pointed out that speaking item types with a recording function are a novel task for kindergarten students.

Interestingly, although a description of the buttons was given through the video tutorial, it appears that the students did not remember this explanation. The interviewers noted that some students (30%) did not pay attention to the tutorial video as the tutorial content was perhaps more accessible to the upper grade students. However, once the students were taught the use of the recording buttons, they were able to use them independently for the rest of the speaking items. Some students enjoyed replaying their responses in order to hear the recordings and sometimes asked the interviewers to re-record their responses.

The interviewers’ comments during the debriefing meetings and their notes indicated that almost all students were able to use the TE features (e.g., drag & drop, zone selection) independently toward the end of the test. This finding suggests that the students at this young grade level are able to quickly learn the use of technology features. Importantly, the finding also suggests that the practice of including sample items will be beneficial for students. The ratings for students’ engagement showed that the student participants (60%) were highly engaged in the given items throughout the testing. However, the interviewers noted that some students (40%) appeared distracted due to their short attention span or, in some cases, very limited English language proficiency, requiring the interviewers’ assistance. As previously mentioned, the interviewers demonstrated what to do in a given item for the very beginning-level students.

***Use of the accessibility tools*.** Among the sample without disabilities, only one non-ELL student used the accessibility tools, namely the magnifier and text-to-speech tools. The student said that he remembered seeing them in the tutorial video. Although the interviewer noted that the student did not require the use of the tools, it was interesting to find that the student learned how to use them through the video tutorial and found the icons to use those tools. The findings regarding students with disabilities are presented in a later section in more detail.

**Grade 1**

Table 6 summarizes the interviewers’ overall ratings in the five major categories on a 3-point scale across the item types. The sample included a total of 15 students without disabilities (11 ELLs and 4 non-ELLs). Two reclassified (i.e., former) ELL students were categorized into the non-ELL group in the analysis. Overall, a similar trend was observed in the Grade 1 sample in which many students (80%) had some difficulties in completing the given item types independently. These students included both ELL and non-ELL students. The interviewers commented that they had to intervene during the testing to provide help with explaining the tasks or the use of TE features. Two factors which appeared to contribute to these difficulties included students’ English language proficiency and their confusion about some of the item directions, as explained further below.

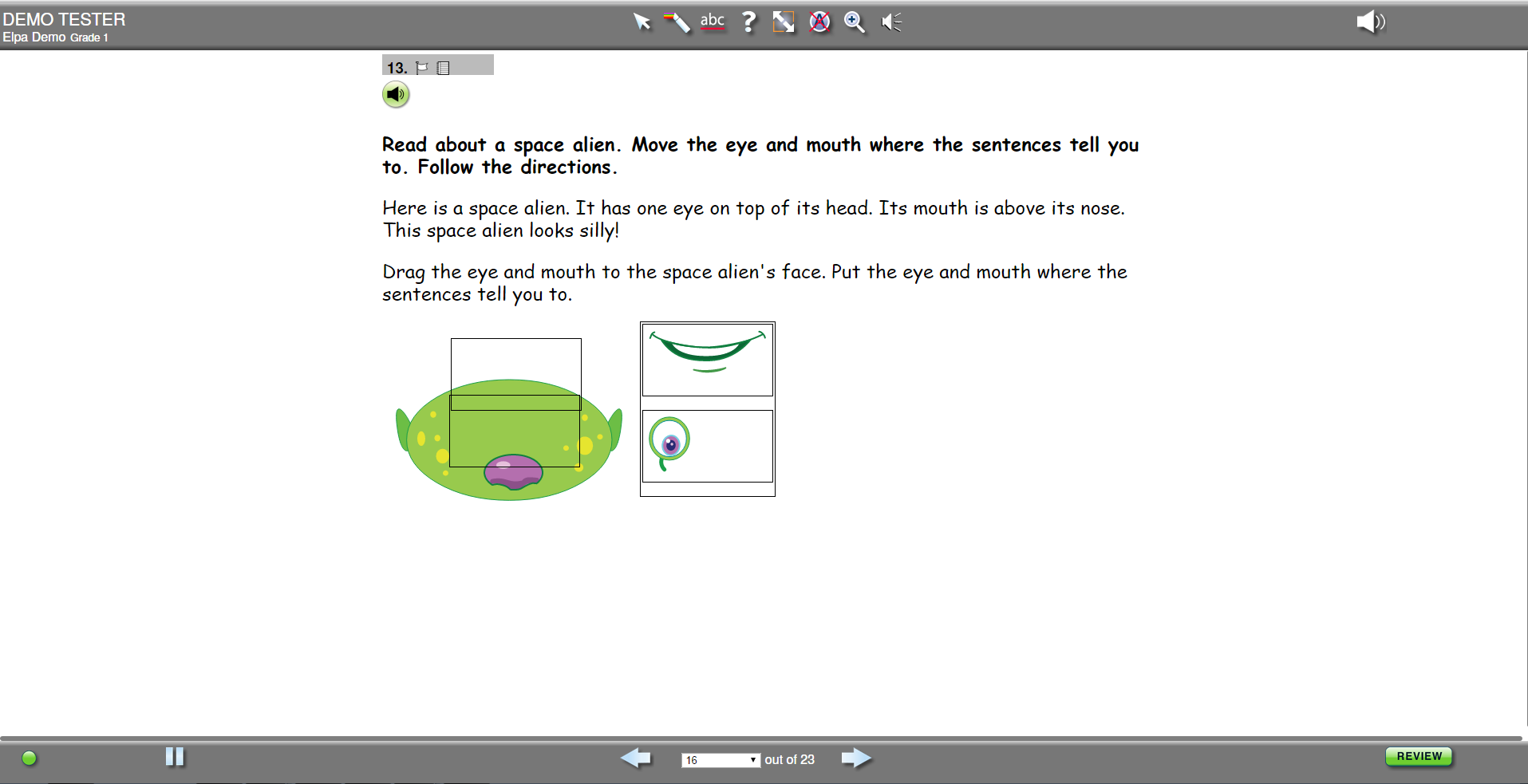
Table 6. The Frequencies of the Overall Ratings across Item Types: Grade 1

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Freq. (%)* | | |
| Categories | 0  no | 1  partially | 2  entirely |
| Understanding the directions | 0 (0) | 11 (73) | 4 (27) |
| Difficulties in the use of TE features | 6 (40) | 9 (60) | 0 (0) |
| Completing the item independently | 0 (0) | 12 (80) | 3 (20) |
| Engagement in the items | 0 (0) | 5 (33) | 10 (67) |
| Using the accessibility tools | 12 (80) | 3 (20) | 0 (0) |

***Clarity of the directions.*** As shown in Table 6, many students (73%) had some difficulty in understanding the directions across the items, on average. Similar reasons to those seen with the Grade K students appeared to contribute to this rating. In particular, the interviewers widely noted that many students had some issues with the directions for the first item of the test, the first speaking item, and some reading items (especially for the “Reading for Details” item). A lack of explicit directions about the arrow buttons to advance to the second item was raised as an issue again. However, once the students were told by the interviewers, this issue was quickly resolved. Although some students were able to use the recording buttons immediately in Grade 1, many students needed to be reminded to press the recording button by the interviewers for the first speaking item. Explicit directions to press the start and stop buttons and/or the opportunity to practice recording seem essential for students, given the novelty of this item type.

Similarly, there were some students who exhibited difficulties in distinguishing the click vs. drag & drop in a few items (e.g., the “Picture Description” item in speaking, the “Word Builder” item in writing). Once directed by the interviewers, the students were able to figure out the rest of the items independently.

Another notable issue emerged from a combination of the direction clarity, item presentation, and students’ reading proficiency in a specific reading item of the “Reading for Details” item. To illustrate this interaction between the test presentation and student characteristics, Figure 2 includes a screenshot of this reading item. First, it was unclear to some students that there was a reading passage that they needed to read in order to complete the item. Some students were confused that only the directions were read aloud and tried to complete the item without reading the given passage. The interviewers commented that the item presentation could be improved to clearly mark the reading passage section in addition to making the directions more pronounced. As is often the case at the beginning of the school year, some students at this grade level had yet to develop their reading skills regardless of their ELL vs. non-ELL status. Thus, those students could not distinguish which text was the passage to read in order to complete this particular item.



*Figure 2.* A sample “Reading for Detail” type item*.*

***Use of TE features in new item types.*** In the Grade 1 sample, all students mentioned having prior experience using a computer except for one ELL student. While 60% of the students were still rated to have some difficulties in using the TE features (e.g., drag & drop, recording buttons, zone selection), only 20% (3 students) were noted to have some drag & drop problems initially. It appears that the students’ limited experience using a mouse was a major source of the drag & drop problem. Related to these issues of using a mouse (e.g., double-clicking, moving, right-clicking), some other students had difficulties in clicking the arrow buttons to navigate the items. The interviewers regularly commented that the arrows seemed too small for these young students who were not precise in their clicks of the mouse and had difficulties with a relatively small target size.

A similar problem was noted in the “Literary Set” item type in reading. This item type included a relatively long passage, requiring the test takers to scroll down to view the entire text. Some interviewers considered this to be a TE feature issue and assigned a score point of 1 to the students who had difficulty with using a mouse and grabbing the scroll bar. All of the interviewers noted that the scroll bar seemed too narrow and the color was too discreet for these young students.

The initial difficulty of using the recording buttons for speaking items was observed again in Grade 1. However, it was also notable that a few students were able to use the recording buttons independently without any explanation from the interviewers, contrary to the Grade K sample. The students also quickly learned how to use the recording buttons. All students were able to complete the second speaking item independently.

Despite the difficulties noted above, students were highly engaged with the items. Some students remarked, “I liked moving the letters. I want more!,” “I like to press the buttons,” “I like doing these activities,” “I want to do more. Can I start again?”  In cases where a score of 1 (partial engagement) was given, it was due to the fact that some students were unable to read fluently yet and thus unable to engage in the reading items.

***Use of the accessibility tools*.** A higher percentage of the Grade 1 sample (30%) were able to use some of the accessibility tools independently compared to the Grade K sample. These students were a mix of ELL and non-ELL students. The students were able to appropriately select and use the tools in certain items. The tools included a highlighter, a line reader, a magnifier, and zoom (up to 4X). One student said that she learned how to use them from “the girl in the video,” suggesting that she understood the tutorial video content. While no students used a magnifier or zoom tool in the “Word Builder” item, a couple of students also commented that the font was too small for them. An interactive demo of the accessibility tools would be one solution to address these students’ comments.

**Grades 2-3**

Table 7 summarizes the interviewers’ overall ratings in the five major categories on a 3-point scale across the item types. The sample in this analysis included a total of 22 students without disabilities (17 ELLs and 5 non-ELLs). All students in this sample had experience with using a computer. A slightly increased percentage (23%) of the sample was rated as being able to complete the items independently without the intervention of the interviewers. Yet, a substantial percentage of the sample (68%), including both ELL and non-ELL students, was rated at 1, or partially able to complete the items independently. As explained below, several possible reasons for this were identified pertaining to direction clarity and recording button issues.

Table 7. The Frequencies of the Overall Ratings across Item Types: Grades 2-3

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Freq. (%)* | | |
| Categories | 0  no | 1  partially | 2  entirely |
| Understanding the directions | 2 (9) | 15 (68) | 5 (23) |
| Difficulties in the use of TE features | 16 (73) | 6 (27) | 0 (0) |
| Completing the item independently | 2 (9) | 15 (68) | 5 (23) |
| Engagement in the items | 0 (0) | 7 (32) | 15 (68) |
| Using the accessibility tools | 19 (86) | 3 (14) | 0 (0) |

***Clarity of the directions.*** As shown in Table 7, a consistent pattern emerged in that the majority of students displayed some difficulties in understanding the directions. The interviewers’ notes were consistent, commenting that the difficulties stemmed from a lack of explicit direction on using the arrow buttons to progress to the second item and the recording buttons to respond to the speaking items. The two students who received a score of 0 in understanding the directions were very beginning-level ELL students, one of whom was a newcomer. The interviewers had to explain the directions for each item to these students.

One notable issue with understanding the directions was observed in the “Picture Caption” item in writing. Some students had problems in completely understanding the directions due to problems with the item presentation. As presented in Figure 3, the directions for the item were not entirely shown in the split screen. Some students were not able to open the left-hand pane and missed the directions stating to use the word bank. This led the interviewers to rate these items as partial understanding of the directions.

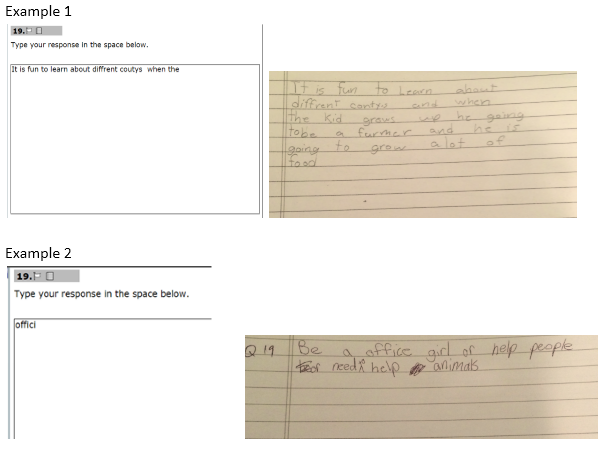


*Figure 3.* A screenshot of a sample “Picture Caption” item in writing.

***Use of TE features in new item types.*** As presented in Table 7, there was a significantly increased percentage of the sample in Grades 2-3 in the “no” difficulty rating for the use of TE features (73%) as compared to Grades K and 1. The students in this grade level also demonstrated easy use of the audio buttons to replay the given listening stimuli in the listening items. The difficulties in using the TE features mainly concerned the recording buttons, similar to the other grade samples. Notably, the students in this grade band had fewer issues with the drag & drop feature. One interesting observation was that two students switched to using a track pad instead of a mouse during testing as they were more familiar with and accustomed to using track pads in their school. The use of a track pad made the drag & drop a little harder for these students compared to the students using a mouse.

Since the test for Grades 2-3 included writing item types that require the students to type on a computer, the interviewers’ notes on students’ typing skills and students’ responses were analyzed. In this study, two sample items from the “Picture Caption” and “Write an Opinion” item types were included. From the study sample, all students except for one Grade 3 ELL student with an advanced ELP level were noted to have substantial difficulties in typing to respond to the items. It was clear that the students in the sample were not accustomed to typing. The students often asked for the interviewer’s help to locate the letters on the keyboard. Many students were also noted to type using one or two fingers at a very slow pace. The students’ response samples provided some evidence that students were not able to fully demonstrate their writing skills. For example, one advanced-level ELL student wanted to type, “I want to be a police officer” for the “Write an Opinion” item type. However, her written response on the computer was, “i wot to be a ples ofi sr.” It appeared that the problem was not only her literacy skills but also her difficulty in finding the letters on the keyboard.

Some students were offered the opportunity to write their responses on paper after spending some time typing on the computer. Figure 4 includes two examples comparing the students’ responses on the computer and paper. It was apparent that the students’ typing skills impeded them in demonstrating their writing skills.



*Figure 4.* Examples of students’ responses on the computer vs. on paper.

***Use of the accessibility tools*.** As shown in Table 7, 14% (3 students) of the sample independently used some accessibility tools. The students included both ELL and non-ELL students without disabilities. The tools they used included a highlighter, a cross-off button, a line reader, a magnifier, a text reader, and zoom.

**Grades 6-8**

Table 8 presents the interviewers’ overall ratings in the five major categories on a 3-point scale across the item types. The sample included a total of 20 students without disabilities in this analysis (17 ELLs and 3 reclassified ELLs). All students in this sample had experience with using a computer. A notably increased percentage (55%) of the sample was rated as “entirely” able to complete the items independently. The interviewers’ notes suggested that their ratings for the partial independence of this sample had to do with the issue of direction clarity, as explained below.

Table 8. The Frequencies of the Overall Ratings across Item Types: Grades 6-8

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Freq. (%)* | | |
| Categories | 0  no | 1  partially | 2  entirely |
| Understanding the directions | 0 (0) | 9 (45) | 11 (55) |
| Difficulties in the use of TE features | 16 (80) | 4 (20) | 0 (0) |
| Completing the item independently | 0 (0) | 9 (45) | 11 (55) |
| Engagement in the items | 0 (0) | 3 (15) | 17 (85) |
| Using the accessibility tools | 16 (80) | 4 (20) | 0 (0) |

***Clarity of the directions.*** The same pattern was observed as in other grades in terms of the ratings for directions. As Table 8 shows, 45% of the sample had some difficulties understanding the directions. As in other grade levels, the students’ initial difficulties were related to a lack of explicit directions for the arrow buttons and recording buttons.

The listening items in the “Listen for Information” and “Interactive Student Presentation” item types contained relatively long listening stimuli. While some students were able to use the audio buttons to replay the stimuli to listen again, others were unaware of the functions of the audio buttons. This finding indicates the need for explicit directions to demonstrate all the functions necessary to complete the items. This is also important to ensure a standardized administration procedure.

Some students were unclear as to whether the given items were reading or speaking items due to the nature of the integrated language skills in the items (e.g., the “Language Arts Presentation” item type in reading). This was another example to indicate the importance of explicit directions.

***Use of TE features in new item types.*** As seen in Table 8, only 20% (4 students) of the sample was rated to have some difficulties in the use of TE features. Several difficulty ratings were related to the fact that some interviewers had to explain the use of the recording buttons to a few students. In addition, one student was observed to have difficulty in using the scroll bar because of its narrowness.

However, the interviewers generally commented that they found a noticeable difference between the Grades K-3 and the Grades 6-8 samples in terms of the use of TE features. None of the students in this sample had a problem with the drag & drop, zone selection, or drop-down menu features. In fact, students were skillful in using the TE features. For the speaking items, they replayed their recorded responses to evaluate their own responses and re-recorded their answers if desired.

The sample of Grades 6-8 students in this study had very little difficulty in typing on the computer. Although only 13 students were able to complete the writing items due to the limited testing time, all students except for one typed their responses in the “Storyboard” item without any typing issues. One student commented that she preferred to write on paper, but other students commented that they preferred to type on the computer due to the neatness of their written responses.

***Use of the accessibility tools*.** In this sample, 20% (4 students) independently used some accessibility tools including a flag for review, a highlighter, a cross-off button, a line reader, a magnifier, a text-reader, and zoom. One student said that he sometimes used a piece of paper or a ruler as a line reader when reading some books and liked the line reader during this test. Besides the accessibility tools, two students asked for a piece of paper for the listening items due to their long listening stimuli. The students took some notes while listening.

**Accessibility and ELLs with Disabilities**

In this section, we present the findings on the 14 ELSWDs across the grade levels in detail (see Table 2 for a summary of the ELSWD sample). Major findings for all ELSWDs are similar to the themes that emerged for students without disabilities in the previous section.

Due to the small sample size for ELSWDs in each grade level, the analysis was qualitative and case-by-case without computing the average frequencies of the interviewers’ ratings. The results for ELSWDs are provided according to three themes across each grade level: clarity of directions, use of TE features in new item types, and use of the accessibility tools.

**Grade K ELLs with Disabilities**

Overall, similar themes emerged from ELSWDs in Kindergarten. The data suggests that all four students generally needed more explicit directions, including some warning or notification that the domain was changing (e.g., listening to speaking, speaking to reading, and reading to writing). Most had some difficulties with using a mouse, and all ELSWDs in Grade K had some difficulty with capturing spoken responses. Most ELSWDs in Grade K also had difficulty reading. These highlights are explained in more detail below.

***Clarity of directions.*** Most students had some difficulty understanding the initial directions. The first item type was to “Follow Instructions” and drag & drop a series of items onto selected target areas of the screen. One student with speech, learning, and behavior difficulties said: *“I gotta click on the robot! But how I click it though?”* After completing the first item, all students had difficulty in advancing to the next item and had to be shown how by the interviewer. This was likely due to a lack of explicit instructions, as discussed in the previous section.

This confusion with directions continued when students progressed to different item types. Some of these transitional difficulties were noted during the transitions from zone items (graphics based) to speaking items, or traditional multiple choice items. For example, when transitioning to the “Picture Description” items, most students (all with speech disabilities and one with multiple disabilities) had some trouble understanding the directions and would click on the pictures instead of or in addition to recording their response with the microphone. However, after additional instruction, most ELSWDs were able to produce spoken responses for this item.

One student with a speech disability also experienced some difficulties in the writing domain. One particular item required students to complete a blank sentence by correctly sequencing the words provided using drag & drop. The student did not understand that multiple words were required to fill in the blank spaces and complete the sentence. He chose only one word as a response and treated the remaining words as distractors, therefore leaving the sentence incomplete. In these instances, some practice and opportunity to gain familiarity with this item type may be helpful for students in future testing situations.

***Use of TE features in new item types.*** Overall, the embedded microphone buttons and some students’ propensity to double click caused some difficulties in capturing their own spoken responses. Some students did not speak during the speaking items, while others had difficulties using the equipment to record a response. One student with a speech and social disability stopped testing at this time. Another student with a speech disability did not record a spoken response with the microphone. When prompted, he demonstrated that he knew how to use the buttons to start, stop, and listen to the microphone recording, but he would not speak.

One of the items in the reading domain (i.e., “Word Wall”) seemed to be somewhat challenging for the two students who reached this point during testing. This item was not rendering on the screen the way the assessment developers intended (the draggers were stretched out beyond the visible screen requiring students to simultaneously scroll and drag & drop). This change from the desired presentation may have resulted in some unintended difficulties for the students, especially since Grade K ELSWDs were already noted to have difficulties with mouse skills.

Additionally, during the reading domain, two students with speech disabilities were observed to have significant difficulty with understanding the split screen display and where the response should be recorded. They continued to click pictures on the left side of the screen, instead of the right side of the screen.

Also noticed during the students’ interaction with items in this domain was that all Grade K ELSWDs had some difficulty with the mouse, including clicking and dragging & dropping. Most students needed hand over hand support. One student with a speech and social disability used two hands to hold the mouse. Another student with multiple disabilities would often forget to pick up the mouse to reposition it. He would run out of room at the edge of the table and say *“I got stuck! I can’t do it!”* The individual attention and assistance was helpful to support this student’s difficulties with using the mouse on the drag & drop items. One student with a speech disability had significant usability issues associated with the mouse (i.e., his clicking) plus interaction with the platform (e.g., small targets; it was very difficult for him to place the mouse in the right spot and click the mouse buttons without exerting pressure on the mouse and moving the mouse off the target). The student also kept pressing the mouse in the middle of the two keys, causing it to lock up (e.g., activating the right click menu). Despite explicit instruction, this pattern was noticed throughout the cognitive lab. This student greatly benefitted from the one-on-one administration, where the interviewer provided consistent hand over hand support to help the student navigate the screen with the mouse.

***Use of the accessibility tools.*** None of the Grade K ELSWDs used the accessibility tools. This could be for a variety of reasons. Although the students saw the tutorial video, the instruction may not have included enough repetition for them to have sufficiently learned about the accessibility tools. Another reason could be that given the students’ young age and the early stage of the school year, it is possible that they had not yet learned in their classroom what those tools are and how they can help support learning.

**Grade 1 ELL with Disabilities**

There was only one ELSWD student in Grade 1, and she possessed a combination of disabilities (i.e., speech, social, reading, writing, math) that may have influenced her performance during the cognitive lab. Additionally, she was quite shy during the interview, even when the interviewer switched into Spanish in an attempt to elicit some dialogue. Further discussion with her teacher confirmed that she is often very quiet and refuses to speak, so her performance during the cognitive lab was not out of the ordinary.

***Clarity of directions.*** The student was observed to consistently skip all textual directions across domains. As a result, it cannot be determined if she thought the directions were clear and easy to understand, or if there were elements that were confusing. She also skipped all of the reading passages. When prompted to try to read, the student explained, *“no sabo”* (translation: I don’t know how), and she moved on to guess at the items.

***Use of TE features in new item types.*** The student recorded three short responses in the speaking section, but then refused to say anything else during the remainder of the speaking domain. She needed persistent prompting from the interviewer to speak for the items. Although the student was noted to have some difficulty with recording her responses, there was not enough evidence to confirm if the performance was due to the microphone buttons, her shyness or other personality characteristics (she was also listed as having a social disability among others in her IEP), or if that was truly the extent of her spoken responses.

***Use of the accessibility tools.*** This student did not use any accessibility tools. Overall, the student’s performance may be tied to her multiple disability diagnoses or her refusal to read the items or directions, or possibly a combination of both.

**Grades 2-3 ELLs with Disabilities**

Overall, the three ELSWDs in Grades 2-3 were better able to interact with the mouse than the ELSWDs in Grades K-1. One student with multiple disabilities including speech, OT [motor difficulties], reading, writing, and math was noted to have a propensity to skip items, although he showed good control of the mouse for logging responses. Almost all of the ELSWDs needed to listen to the stimuli more than once, and they benefitted from more directions including to help them transition across domains. As noted in the highlights below, in instances where students had difficulties, they were also able to quickly learn and benefit from additional instruction from the interviewer. These instances clearly point to the importance of additional opportunities to practice for ELSWDs.

***Clarity of directions.*** Throughout the listening domain, most students had to listen to the initial directions more than once to understand what they needed to do. Most were able to work somewhat independently and used the green buttons easily to replay the directions.

***Use of TE features in new item types.*** When the students interacted with the items in the speaking domain, they seemed to have some difficulty recording complete responses. For example, although all three students had difficulty understanding the directions and needed to listen to the directions more than once, two students with multiple disabilities had difficulty understanding how and where to record the answer. Instead of using the microphone, they were trying to click on the circled letters at the bottom of the pictures and could not record their response.

The split screen was also somewhat confusing for the screen presentation layout (i.e., reading text on the left pane and answering questions on the right pane) for one student in particular, the student with multiple disabilities including motor difficulties. Perhaps additional opportunities to gain familiarity with the split screen layout would help test takers with a similar background or response pattern in future instances.

During the writing domain, the three ELSWDs demonstrated some difficulties in typing their responses. One student (who had social and learning difficulties) struggled to begin typing and asked the interviewer how to spell the first word of the sentence (i.e., “they”), while others (with similar disability combinations including motor difficulties) used the “hunt and peck” method with one finger to type. One student with multiple disabilities had accidentally selected “Caps Lock” but did not unselect it so part of her response is in caps: *“the boy is play throw bALL.”* Her next response was recorded on paper, and she wrote *“I boy like to be a Farmers.”* Overall, while students were able to type a response, some were very inefficient at doing so. Typing, then, may be a skill these students had not yet had the opportunity to learn, making the scratch paper option (available during the cognitive lab session) an important tool for students who needed to use it.

***Use of the accessibility tools.*** As summarized in Table 9, all three students in Grades 2-3 used some accessibility tools. However, one student with multiple disabilities, including motor difficulties, seemed distracted by the accessibility toolbar and was observed hovering over the tools or clicking on them during the stimulus presentation.

Table 9.The Use of Accessibility Tools, Grades 2-3 ELSWDs

|  |  |  |
| --- | --- | --- |
| Disability Categories | ELP Levels | Accessibility Tools Used |
| Reading, writing, math, social | Advanced | Text reader |
| Reading, writing, math, social | Advanced | Line reader (when prompted) |
| Speech, OT (motor), reading, writing, math | Advanced | Zoom; help button; highlighter on speaking item; magnification on writing item |

Overall, some students did experience some difficulties that were ameliorated with the use of the accessibility tools or by the use of the embedded supports. For example, one student used the text reader to help read the content on the screen, while another student used a line reader when prompted. Moreover, both students had difficulty understanding the directions and needed to hear them more than once.

**Grades 6-8 ELLs with Disabilities**

Overall, most ELSWDs in Grades 6-8 were able to interact more independently with the platform and content compared to the younger students in Grades K-3. Most students in this grade band used the accessibility features available to them, and some ELSWDs demonstrated strategic use of these tools. Also unique to Grades 6-8 was the amount of reading required. As observed during the cognitive labs, students took a long time to read a passage in general. Specific findings are explained in more detail below.

***Clarity of directions*.** During the listening domain, some students had difficulty with the audio and directions. One student with learning disabilities had some difficulty interacting with the audio buttons on the platform to find the audio she wanted to play again: *“Which one is it? There are 5. That’s confusing.”* Another student diagnosed as having a disability under the category of “other health impairment” had some difficulty finding the written directions on the screen and asked the interviewer to explain how to record her answer. When prompted to try, the student dragged and dropped successfully; similarly another student with a learning disability was observed initially as hesitating but was then able to respond correctly.Other students(one with a learning disability, the other with a visual impairment) had some difficulty with the directions and tried to click & click rather than drag & drop. Similarly, another student who had a learning disability had some difficulty following directions and knowing that he needed to select two responses as stated in the directions.

***Use of TE features in new item types*.** Multiple difficulties were noted with the TE features for Grades 6-8 ELSWDs. Two students (one with a learning disability, the other with a visual impairment) had some difficulty understanding where to record the response for zone items. However, one student with a learning disability was observed using the color that shows up when the mouse is over the zone item as a strategy to learn where to record her response.

Another student with a learning disability had some difficulty understanding the split screen feature. This split screen, coupled with her difficulty understanding the directions, led to her asking for clarification about the graph she was supposed to talk about, the one on the left (the stimulus) or the one on the right (the recording graphic next to the microphone buttons). The student was interpreting the sound wave graphic next to the microphone buttons as another graph on the page, and thus was not sure which graphic to talk about.

Some students displayed difficulty in understanding how to accurately respond to some of the reading items. One student with a learning disability did not understand the functionality of the underlined word, and she tried to search for the target word manually through the reading passage. When prompted to click on “guarded” in the stem, she exclaimed, “*It took me to the bottom! Oh!*” She immediately understood the intended functionality after the interviewer’s explicit instructions. For writing items, all three ELSWDs (with learning disabilities) displayed no difficulty in typing a response.

***Use of the accessibility tools.*** Table 10 presents a summary of the use of accessibility tools by the students with specific disabilities.

Table 10. The Use of Accessibility Tools, Grades 6-8 ELSWDs

|  |  |  |
| --- | --- | --- |
| Disability Categories | ELP Levels | Accessibility Tools Used |
| Visual impairment (late onset) | Advanced | Zoom, magnification; attempted text reader |
| Specific learning disability | Advanced | None |
| Other health impairment | Intermediate | Zoom |
| Specific learning disability | Intermediate | Highlighter; attempted text reader |
| Specific learning disability | Advanced | Highlighter |
| Specific learning disability | Advanced | Highlighter; arrow; opened the digital note pad but did not use |

Students were also observed demonstrating various strategies to interact with items in the speaking domain. One student with a learning disability used the highlighting tool to highlight the title, author, setting, main characters and parts of the remainder of the stimulus during the read aloud; he also took time after the audio read aloud to look at the text and highlight select phrases. Two students with learning disabilities also used the highlighter tool to highlight the directions. One of these same students strategically used the audio buttons to replay the stimulus to hear the author’s name, and after the author’s name was said, the student shut off the audio and recorded his response.

Although a significant amount of time was allocated for students to read the passages, some students read more efficiently than others, while others were more reluctant. One student with other health impairment had no trouble using the scroll bar to read the text but noted that it was too long for her liking. Another student with a learning disability had some difficulty with decoding irrespective of the use of the line reader. For example, when prompted about “guarded” she read it as “gargle” and laughed to herself, acting it out (gargling with the throat).

One student had a late onset visual impairment. She had some unique difficulties interacting with the test platform that seemed to be uniquely associated with her disability. As a result, these accessibility difficulties are reported in detail. She needed help navigating the content due to some difficulty seeing the directional arrows (i.e., target size). When the student used the magnifier to see the arrows, the magnifier would not let her click the arrows underneath. The student realized she would have to close the magnifier to be able to click on the arrows to move on. She also asked the interviewer for confirmation about recording; this may be due to the fact that the student was not able to use the zoom and magnification tool to verify where she was clicking on the screen. Overall, the student was able to record her responses successfully with the use of the accessibility tools.

This student with a visual impairment did have some difficulty viewing the content on the screen with only the magnification or the zoom accessibility tools. This was noticeable for the book report visual in the speaking domain. The interviewer explicitly taught the student how to layer tools (e.g., zoom and magnifier) to see the book report. Without this layering, the student had significant difficulty seeing the information in the book report, even when using the zoom feature. Related to using the zoom feature was her difficulty with using the scroll bar and zoom feature together; first, she would activate the zoom feature, then she would have to scroll left/right/up/down to see the content on the screen (i.e., the larger the zoom, the more scrolling that was required). Unfortunately, the student had some difficulty with the additional scrolling because the scroll bar target did not enlarge with the rest of the screen when she used the zoom feature. Similarly, the embedded buttons for audio replay caused some difficulty due to the small visual cue.

She also could not record a response for the drop down menu item types although she knew how to interact with the drop down menu. The student had difficulty seeing the drop down choices since they did not enlarge with zoom feature (although the rest of the passage did). The student overlapped the zoom and magnifier to make them work together, but she realized again that she couldn’t click through the magnifier to make the menu drop down to select her answer.

Another accessibility difficulty was noticed with the zoom feature in the reading domain. The split screen feature (passage on the left pane, item on the right pane) and the zoom feature, combined with transitioning to the next items, resulted in the screen being focused on the right pane for the following item. As a result, this student did not see the story on the side (i.e., left pane) due to the zoom. She was observed as immediately trying to answer the question. When interviewed, she explained she thought the question was tied to different content and that maybe it was a discrete question and not tied to a reading passage. When prompted that there was a reading passage, she did independently begin using the scroll bars to look for a passage, suggesting that the lack of attention to the reading passage was accidental, not intentional.

**Conclusion and Recommendations**

This study was undertaken to try out a sample of newly-developed items with TE features prior to a large-scale field test. To ensure that the TE items function appropriately without causing any systematic sources of construct-irrelevant variance, it was important to conduct an empirical investigation about students’ interaction with the items during the assessment development stage. Thus, the study focused on students’ abilities to use TE features to demonstrate their English language proficiency and any possible sources of difficulties related to item quality such as the clarity of directions and the presentation of the items.

Overall, the study yielded promising results indicating that the students in the sample were able to learn how to use TE features quickly during testing and were highly engaged in completing the items on a computer. The majority of ELSWDs were also able to successfully interact with the TE features and platform after receiving additional, targeted instruction from the interviewer. A set of clear patterns emerged in the types of difficulties students encountered while completing the items. It is important to note that this study was conducted with a small sample using a preliminary test platform when the directions and final design of the test platform were yet to be developed. Thus, one common source of difficulty was largely due to the fact that the assessment platform was in the process of development. While the study findings need to be carefully interpreted, they provide valuable information to enhance the quality of the test.

Based on these findings, we have attempted to offer the following key recommendations and suggestions for future development or administration of the operational ELPA21 assessments. Each is followed how the recommendation/suggestion was addressed in the field test and or will be addressed post field test[[3]](#footnote-3).

* **A tutorial and sample practice items should be provided to the students prior to testing.** Considering that the ELPA21 assessments will include various, innovative TE item types on a computer-based test platform, it is critical to provide opportunities for the students to become familiarized with the item types including the various TE features and other platform functions. This is essential in order to adequately elicit students’ English language proficiency skills, unimpeded by the use of TE features. We recommend that the tutorial should include accessible language and visuals (including animations) that are appropriate to the specific grade and age group. It should be presented in such a way that ELL students understand the content. It is also equally important that students have hands-on experience with technology features through practice items.
  + ELPA21 provided an interactive item demonstration site for the field test, accompanied by detailed lesson plans, for each of following the grade bands: K, 1, 2-3, 4-5, and combined 6-8 and 9-12. The demonstration items covered each of the item interactions in the grade-level texts. Teachers were encouraged to both demonstrate the items to their students and provide their students with opportunities to practice with the items so they would be familiar with the testing environment.
* **Clear and explicit directions must accompany each item type.** The study results suggested that the students at younger grades (i.e., Grades K-2) were sometimes unclear about the expected response format such as zone selection vs. drag & drop and clicking a choice vs. speaking about the choice. Thus, the directions should clearly include what the students are expected to do in order to complete the given item types. For the students whose English language proficiency is too low to understand the directions, individual demonstrations of expected responses with sample items would be desirable, in addition to prior exposure to sample practice items.
  + Each item type on the field test is accompanied by specific directions. Student responses and reports from field test administrators will be reviewed to determine the effectiveness of the instructions.
* **The construct measured in each item type should be clearly defined in order to determine any necessary accommodations or specific accessibility features for a computer-based test.** This issue was particularly a consideration for some writing items that require students to provide an extended writing sample. In our study, some students at the elementary grades demonstrated a noticeable difference when writing on paper compared to writing on a computer. If the construct of interest in a writing item was to solely measure students’ writing proficiency regardless of the test delivery mode, allowing students to write on paper should be an acceptable response format. That is, an explicit definition of the construct for each item type will help the test users to determine appropriate methods for alternate responses or accommodations to support students in need.
  + The *Accommodations and Accessibility Framework* that guides ELPA21 takes the constructs being assessed into account, and items and tasks were developed with available accommodations and accessibility tools in mind. Student responses and reports from field test administrators will be reviewed to determine whether the administration procedures are appropriate, particularly attending to the Writing issues noted.
* **Accessibility tools designed to support all students’ access to the test content should be designed to work with innovative TE item types and the computer platform; the tools must be demonstrated to all students prior to testing.** In order to fully realize the benefits of technology in assessment, it is evident that students should be provided with the opportunity to practice the use of the accessibility tools. The tools also need to be designed to work with the specific test platform. For example, if the screen is zoomed, the content in the drop-down menu should zoom in along with the rest of the content on the screen. Additionally, the zoom should also enlarge the standard platform features (accessibility tool bar, scroll bar).
  + The interactive item demonstration described above includes the accessibility tools that were used on the field test. Reports from field test administrators will be reviewed to evaluate how well the tools worked.
* **Individual assistance should be provided during the test administration for younger grade students.** The study results indicated that students in Grades K and 1 had limited testing experience and were accustomed to confirming understanding with the teachers. The students also had limited experience with technology. Thus, we recommend that individual administration of the test or individual assistance during the test be provided to these particular grade levels. 
  + The recommended administrator-to-student ratio for Grades K and 1 for the computer-based portion of the assessment is 1 to 5; for Grade 2 it is 1 to 8. Reports from field test administrators will be reviewed to determine whether this ratio is sufficient to allow individual assistance for these students.
* **Individual assistance must be provided for ELL students with disabilities during the testing in addition to ample opportunities to gain familiarity with TE features and the testing platform**. It takes extra support to administer a computer-based ELP test for ELSWDs. The complexities lie in the students’ unique needs, factoring in the varied ELP levels, grade levels, and disabilities. While the study findings suggest that most ELSWDs were able to interact with the TE features, ELSWDs needed individual assistance to demonstrate how to use these TE features and accessibility tools. In our study sample, students with certain disabilities seemed to have unique difficulties. For example, students with speech disabilities in Grades K-1 all had some difficulty with recording features in the speaking items. These students should be provided multiple opportunities to practice the use of new TE features.
  + The recommended administrator-to-student ratio for Grades K and 1 for the computer-based portion of the assessment is one administrator for one to five students, and for the paper-based portion it is one administrator for one to three students; for Grade 2 it is one administrator for one to eight students. Reports from field test administrators will be reviewed to determine whether this ratio is sufficient to allow individual assistance for these students.
* ELSWDs should be provided with individual assistance during test administration, as well as sufficient opportunities to practice using the TE features and platform prior to testing.
  + The interactive item demonstration site provided practice with designated accessibility features and embedded accommodations for students who would use them during the field test. In addition, the Test Administration Manual provides information about what type of help can be provided to students during test administration, such as assistance with scrolling, clicking, and starting/stopping the recordings.
* **Item formatting should be clear and intuitive and should not require unnecessary use of technology navigation features.** It takes extreme care to present new item types on a computer in a way that students easily understand the item layout and have easy access to test content. It is preferable that students can view an entire item without scrolling the screen so that students easily understand what they are expected to do. Yet, presenting a reading passage requiring students to scroll is often unavoidable. In that case, the directions should be clear and the scroll bar should be noticeable. The appropriate size of the font, graphics, and functional icons also need to be carefully determined.
  + Improvements were made to the field test platform based on the cognitive lab results, and more are planned for the operational platform. In addition to the cognitive lab results, reports from field test administrators will be reviewed to determine what changes are needed.
* **Technical requirements should include detailed specifications about the equipment and must be tested prior to testing.** To avoid a lack of standardization of administration, it is critical that the test users be informed of the detailed specifications. For example, the optimum screen size, the desirable size of a mouse, the type of headset with microphone, the operating system, and internet connectivity should be detailed and tested in advance.
  + ELPA21 developed a *Setup and Installation Guide* that includes detailed specifications for hardware and bandwidth, tryout guidance, and tips for making sure the testing system was implemented properly.

As noted earlier, the limitations of the study are evident, including a very small sample size from two urban school districts. It must also be reiterated that the test platform for this cognitive lab study was a preliminary one, differing from a large-scale field test or an operational platform. Thus, the types of difficulties students experienced due to the preliminary platform may not be generalizable and should be interpreted in the given study context.

Despite the limitations, the results of the study offered valuable information. The results also signify the need for further empirical research. One area of research should investigate whether future modifications made to address some of the issues reported on in this study improve the testing practice for students. Sustained research will be needed to monitor the appropriate use of the technology features and provide useful information to enhance the validity and technical qualities for the ELPA21 assessments.

**Acknowledgements**

The authors wish to acknowledge the support for conducting this study from CCSSO and ELPA21, including Cat Still and Phoebe Winter. The authors also acknowledge the study team members, Alexis Lopez, Cindy Nguyen, Emilie Pooler, Lorraine Sova, Florencia Tolentino, and Katie Vlasov at ETS as well as the QAI team. The authors are grateful to Kim Hayes from the Washington State OSPI for her invaluable assistance in recruiting the schools and sharing her helpful insights. Special thanks go to the participating schools, teachers, and students.

**References**

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing.* Washington, DC: Author.

Martiniello, M. (2008). Language and the performance of English language learners in math word problems. *Harvard Educational Review*, 78 (2), 333-368.

Winter, P. C. Kopriva, R. J., Chen, C. & Emick, J. E. (2006). Exploring individual and item factors that affect assessment validity for diverse learners: Results from a large-scale cognitive lab. *Learning and Individual Differences, 16,* 267–276.

Wolf, M. K., Kim, J., & Kao, J. (2012). The effects of glossary and read-aloud accommodations on English language learners’ performance on a mathematics assessment. *Applied Measurement in Education, 25(4)*, 347-374.

**Appendix A**

**Cognitive Lab Interview Questions and Rating Scales**

**COMPUTER QUESTIONS**

1. Have you used a computer before? Yes: \_\_\_\_ No: \_\_\_\_\_\_
2. How often do you use a computer? (1) never, (2) once a month, (3) two to three times a month, (4) at least two to three times a week
3. Did you do any activities on the computer in school? Yes: \_\_\_\_ No:\_\_\_\_\_\_
4. What activity did you do? How often did you do it? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Questions and Prompts**

**Ask the following questions after each set of questions (i.e., each item type). Adjust your language as needed for the student you interview.**

* **Clarity of Directions:** Check student’s understanding of directions
* ***Did you understand the directions? What did you need to do?***
  + *Can you tell me what you were asked to do?*
  + *Were there any parts that you did not understand?*
  + *Were there any words that you did not understand?*
* **Difficulty:** Check any difficulties the student had.
* ***Were these questions easy to answer or hard to answer?***
* *(Very easy, A little bit easy, A little bit hard, Very hard)*
* ***Why do you think it was easy/hard? Which part?***
* *Why did you choose that as an answer?*

**[For speaking items]**

* *Did you have any difficulty in using the headset and microphone? Why?*
* *Did you have any difficulty in recording your response in the speaking question? Why? (i.e., play, pause, stop, listen, replay)*
* **Question Presentation:** Check student’s understanding of how the question is presented.
* ***Did you understand this picture? Can you tell me what this picture is about?***
* ***Was the picture/letter/word big enough for you to understand?***
* *Can you tell me where you look first on the screen? Where do you look next?*
* *Were you able to scroll up and down the screen?*
* *Were there any parts that you did not understand on this screen?*
* **TE Features in Question:** Check student's ability to use TE features.
* ***Were you able to move [the picture/object]? Did you have any difficulty with moving (drag & drop) the picture? Why?*** *[ask the same question for other features such as hotspot and multiple sections]*
* *Was it clear where you needed to click?*

**GENERAL FOLLOW- UP QUESTIONS**

* *What did you mean by \_\_\_\_\_\_?*
* *Can you tell me more?*

**AFTER THE TEST**

* **Usability of Other Technical Functions:** **Record student’s difficulties with interacting with the test question or the testing interface (computer, mouse, keyboard, recording, navigational pane).**
* *Did you have any difficulty in using the mouse? Clicking? Why?*
* *Did you have any difficulty using the scroll bar to see the full screen?*
* *Did you have any difficulty in using the headset and microphone? Why?*
* *Did you have any difficulty in recording your response in the speaking question? Why? (i.e., play, pause, stop, listen, replay)*
* **Any Specific Difficulty:** **Record student’s difficulties with any specific questions.**
* *Which one was most difficult for you? Why?*
* *Were there any words you did not understand?*
* *Were there any parts of the computer that were not easy to use? Give me examples.*
* *Which one was easiest for you? Why?*
* *Which one did you like the most? Why?*
* **Accessibility Features:** **Record student’s difficulties with using the available accessibility features (including universal and designated supports).**

Note: show screen shot of accessibility features tool bar when asking these questions

* *Did you understand what these pictures were? Go over each tool 1x1*
* *Why did/didn’t you use them?*
* *(If the student used the features) Were these helpful? Why/why not?*
* *Was anything confusing or hard to understand?*

**Rating Scale**

The Interviewer completes the following ratings for each item type and overall test performance.

|  |  |  |  |
| --- | --- | --- | --- |
| Topic | Rating & Comments | | |
| Did the student understand what he or she was supposed to do based on the question directions? (**Directions)** | 0 (no) | 1 (partially) | 2 (yes, entirely) |
| Comments: | | |
| Did the student understand the layout of the question? **(Question Presentation)** | 0 (no) | 1 (partially) | 2 (yes, entirely) |
| Comments: | | |
| Were there any issues or difficulties with use of the TE features (e.g., drag and drop, zone selection, hot spot, writing tools—cut, copy, paste)? **(TE Questions)** | 0 (no) | 1 (partially) | 2 (yes, entirely) |
| Comments: | | |
| Did the student have any issues with using the equipment (keyboard, mouse, recording, headphones, microphone)? **(Equipment)** | 0 (no) | 1 (partially) | 2 (yes, entirely) |
| Comments | | |
| Did the student have any issues or difficulties in making use of the accessibility features (universal and designated features)? **(Accessibility- Usability)** | 0 (no) | 1 (partially) | 2 (yes, entirely) |
| Comments | | |
| To what extent was the student engaged? **(Engagement)** | 0 (no) | 1 (partially) | 2 (yes, entirely) |
| Comments | | |
| Was the student able to complete the item(s) independently without difficulties? **(Difficulty/Independence)** | 0 (no) | 1 (partially) | 2 (yes, entirely) |
| Comments | | |

Check the extent to which the student used the given accessibility features (use for ALL students)

|  |  |  |  |
| --- | --- | --- | --- |
| **Accessibility Tools (Features available to all test takers)** | **Overall, was this feature used?** | | |
|  | 0  (No Use) | 1  (Minimal Use) | 2  (Frequent Use) |
| Digital notepad |  |  |  |
| Writing tools: bold, italics, bullets, undo/redo on digital notepad |  |  |  |
| Spell check on digital notepad |  |  |  |
| Flag for review |  |  |  |
| Arrow |  |  |  |
| Highlighter |  |  |  |
| Line Reader |  |  |  |
| Help Button (question mark) |  |  |  |
| Cross off button (answer choice eliminator) |  |  |  |
| Zoom (up to 4X) |  |  |  |
| Magnification device |  |  |  |
| Read aloud directions (text-to-speech/text reader) |  |  |  |

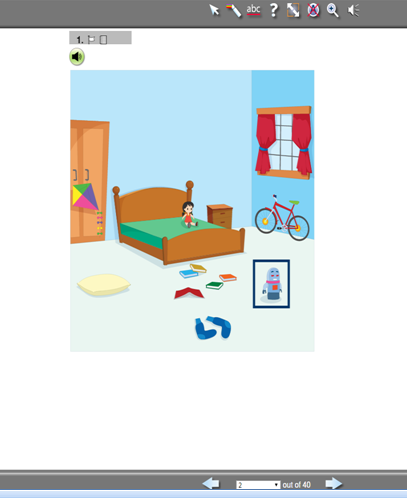
**Appendix B**

**Specific Issues Observed During the Cognitive Labs**

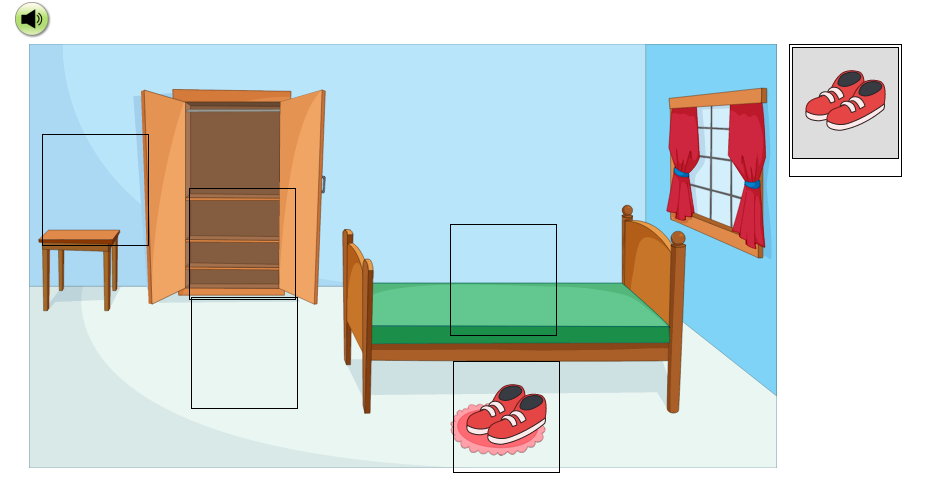
The description below is part of the debriefing document describing specific issues the interviewers noted.

1. **Issues related to directions**

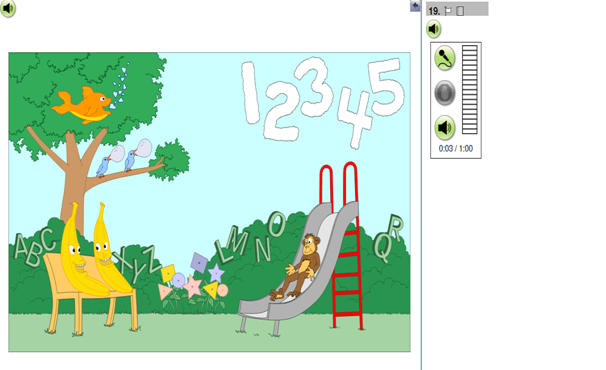
* Almost all students didn’t know what to do upon the completion of the first question due to the lack of directions or design features (e.g., Next button)



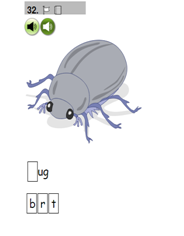
* Some confusion in the following direction items was due to the fact that the picture still appeared in the question box after being moved.



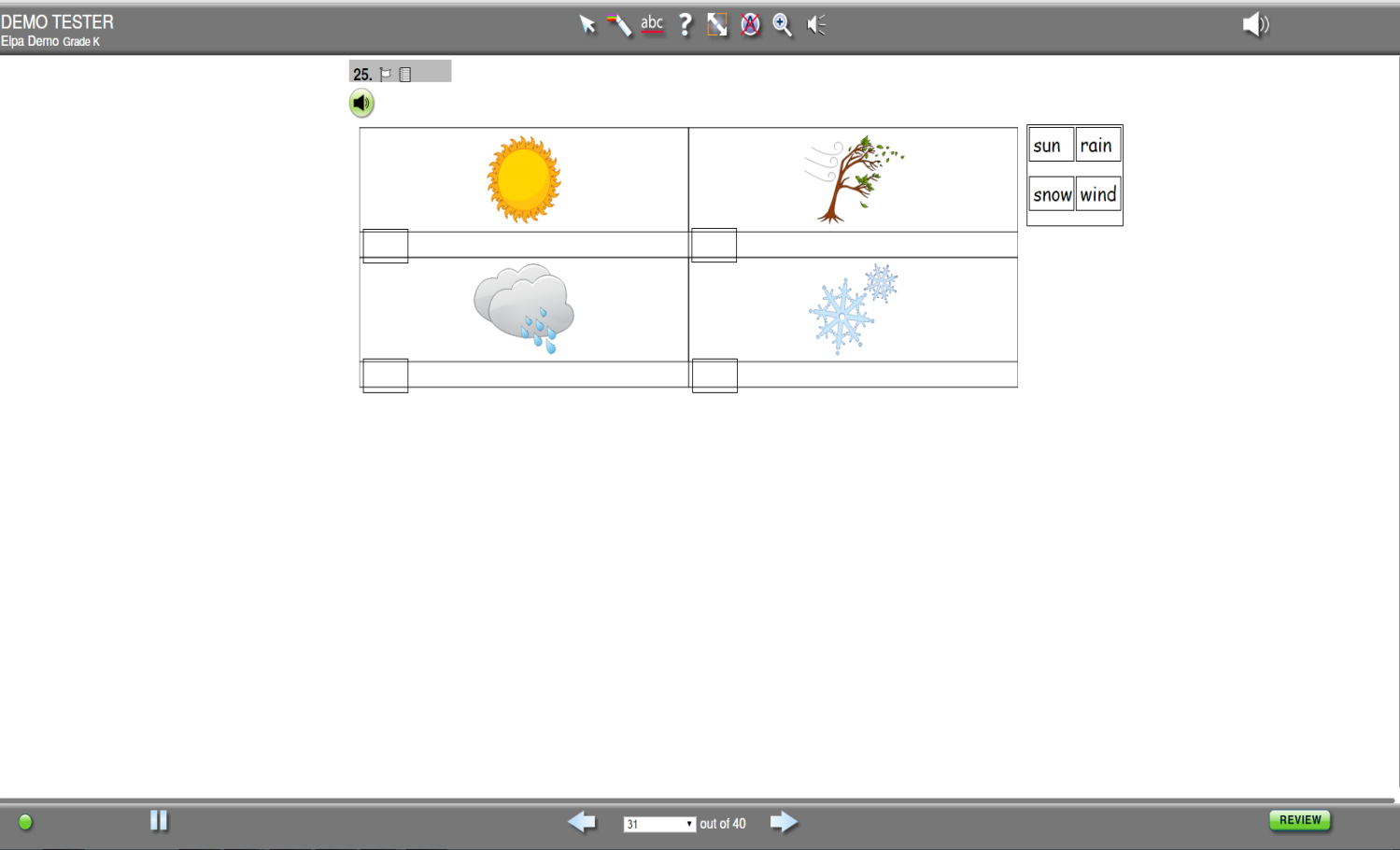
* Some general directions needed to be clearer for each domain. Many students were confused when the domain would change because there was no indication of the domain transition. (e.g., It was unclear for the students when it was a speaking item, since the previous items were in the listening domain, where they had to click on various targets. For the speaking items, including silly picture, the students wanted to click on the silly things in the picture.)



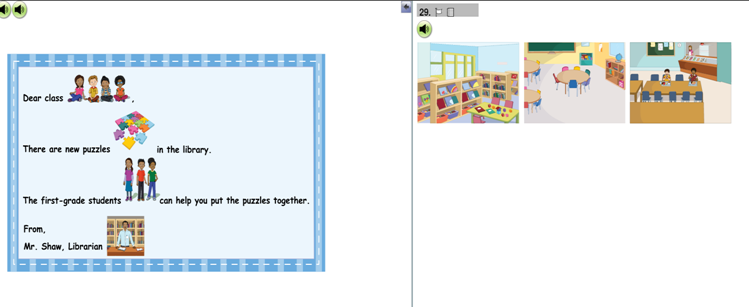
* It was unclear for the students when it was a speaking item. They did not seem to know what to do after they listened to the instructions. They did not realize they needed to record their answers and needed prompting from the interviewer to use the recording buttons. When the interviewer helped them to use the recording, some students still turned to interviewer for confirmation that they used the recording properly. This happened often and supports the need for one-on-one instruction and support, especially for Grades K and 1.
* Particularly for Grades K and 1, some students were confused when to drag & drop targets vs. click targets. For example, students would often click the letter instead of dragging and dropping the letter.



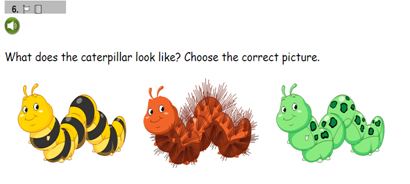
* In a similar example, some students didn’t realize that this is a drag & drop. Some Grade K students had difficulties in drag & drop, due to various issues such as using the mouse and motor control as well as table real estate.



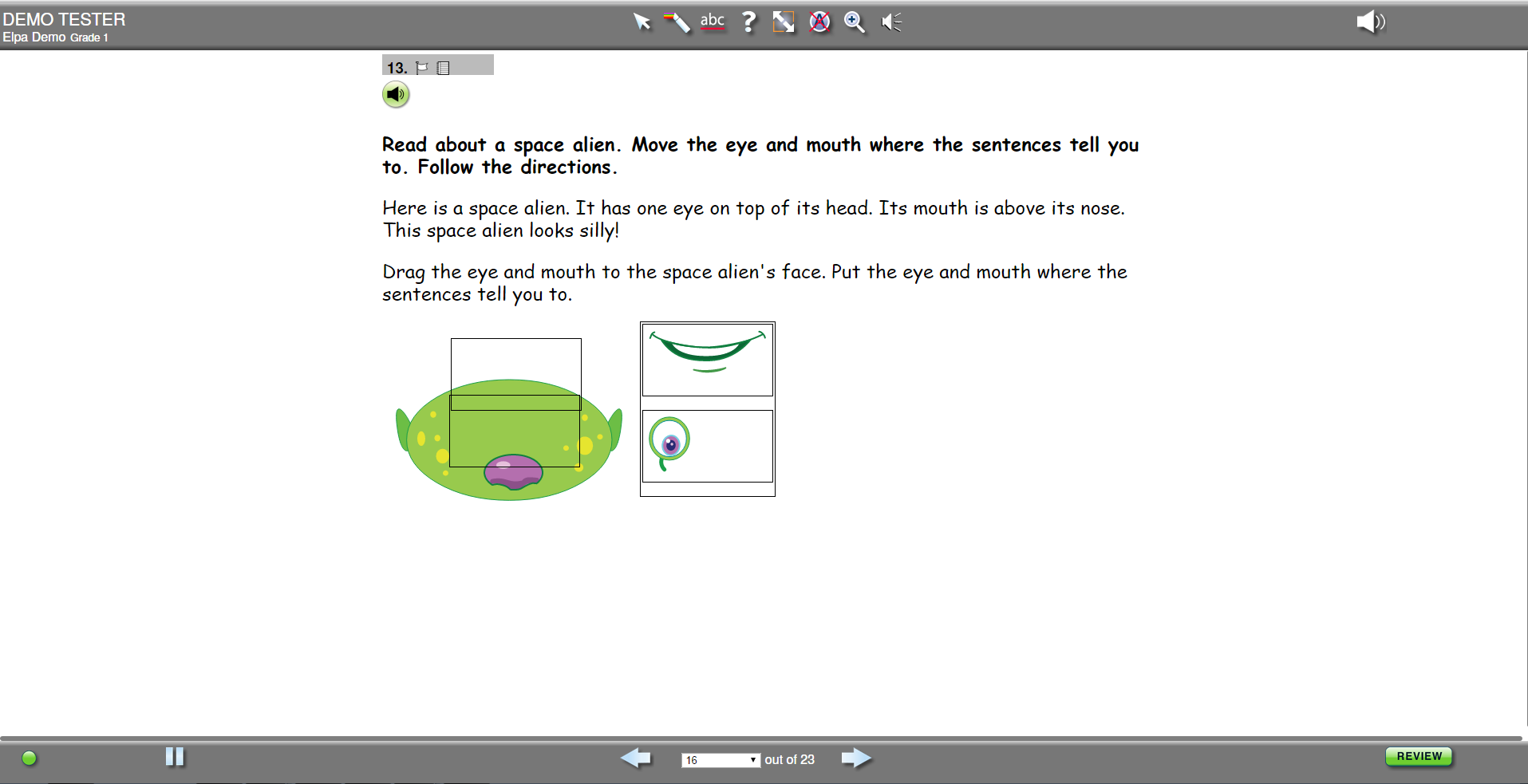
* Some students were confused with where to click to select an answer in the following item due to the pictures in the reading passage (also might be due to the split screen):



* In the listening section, some Grades K and1 students did not seem to understand that they were supposed to answer the questions based on the specific listening passage. Directions could be more explicit or repetitive (or in the admin guide or in the practice items). e.g., caterpillar item

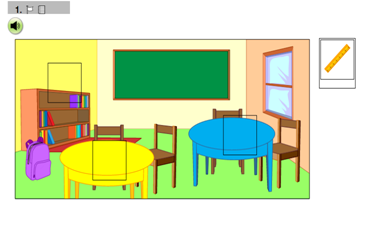


* Some inconsistency in the format caused some confusion (e.g., the space-alien item for Grades 1-3). Many students did not understand that this was a reading item and didn’t understand which sentences to read. Perhaps having a box (or other consistent presentation) around the reading passage might be helpful to indicate that there is reading passage. The overlapping drop zones (boxes) also caused some confusion.



**2. Issues related to the use of TE features**

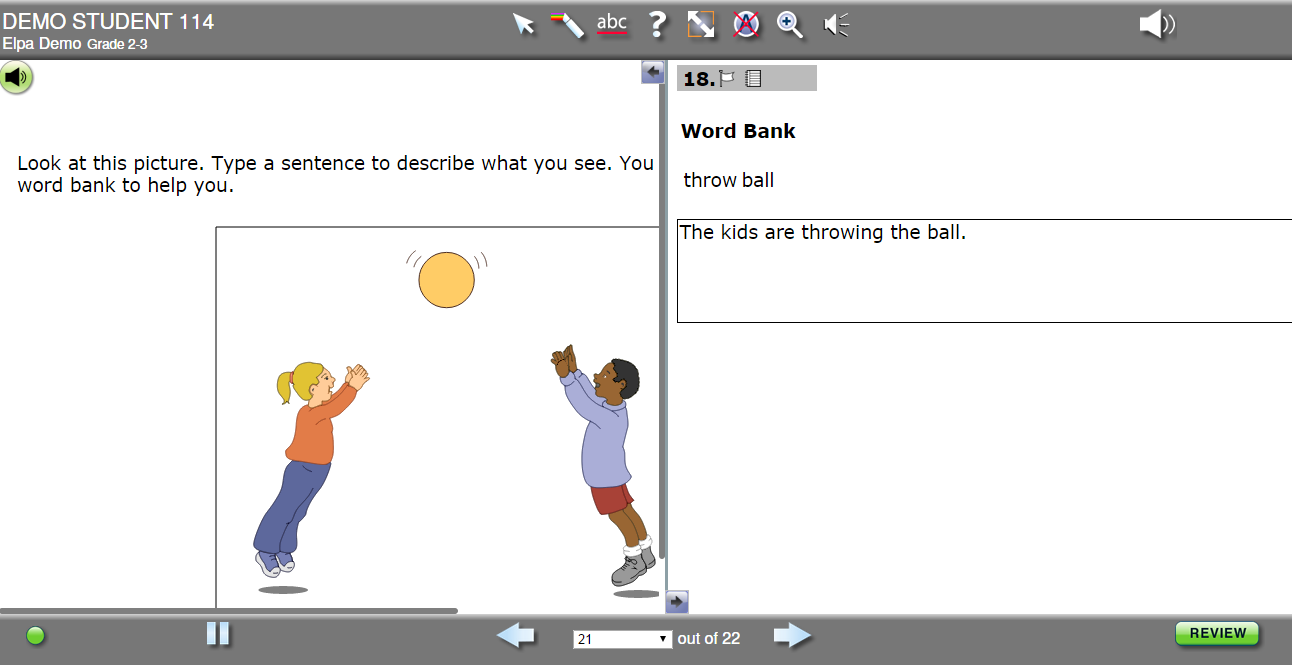
* Drag & drop: mostly no issues, except for Grades K and 1. Some students had trouble with drag & drop, particularly when they needed to drag the picture to a relatively far location. Drag & drop seemed hard for the students who were used to using the track pad of the laptop.

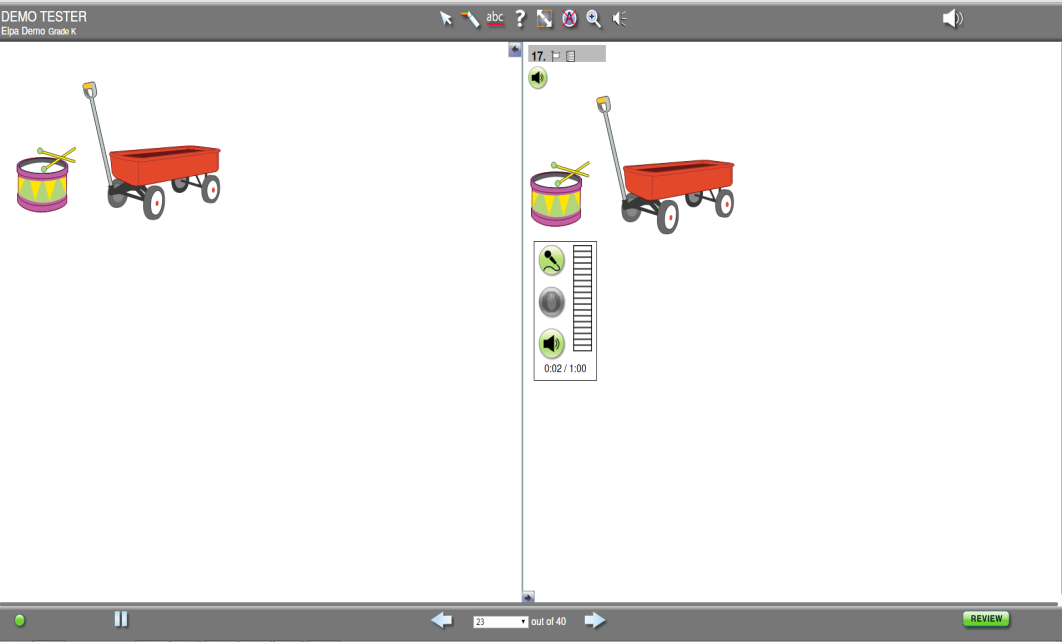


* Zone selection: mostly no issues
* Audio buttons: Grades K-3 students had mixed level of audio button uses. Grades 6-8 students were able to use them better.
* Recording buttons: Almost all students at Grades K-3 needed to practice first with instructions. Some students were confused with when to start and stop recording. Ample practice is needed for Grades K-1.
* Mouse use: Some students preferred to use the track pad instead of the mouse since they are accustomed to using this in school, but this was challenging for the drag and drop options (Grade K). Some students needed help using the mouse (even at Grades 2-3). A smaller-sized mouse for Grade K is desirable.

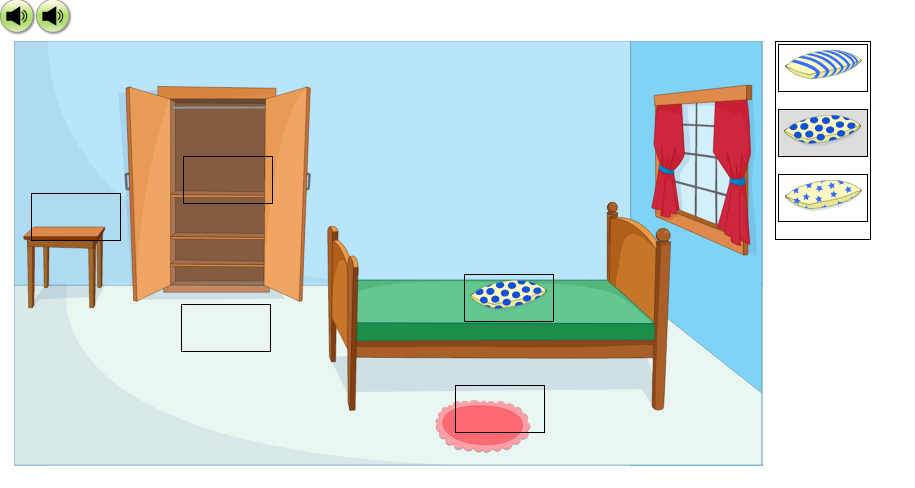
1. **Some technical and platform design issues**

* Laptop or setting specifications: even though all the laptops had the same screen size, one laptop did not show the pictures in one screen, which required additional scrolling. For our cog lab sample at Grades K-3, the students were not accustomed to using the scroll bar. The scroll bar was too narrow for them. For an ELL student with a visual disability, the scroll bar did not enlarge with the zoom feature, making it difficult for her to scroll independently without visual confirmation with the zoom and magnification layering.
* The split screen for Grades K-3 students seemed to cause confusion, requiring additional technology skills.

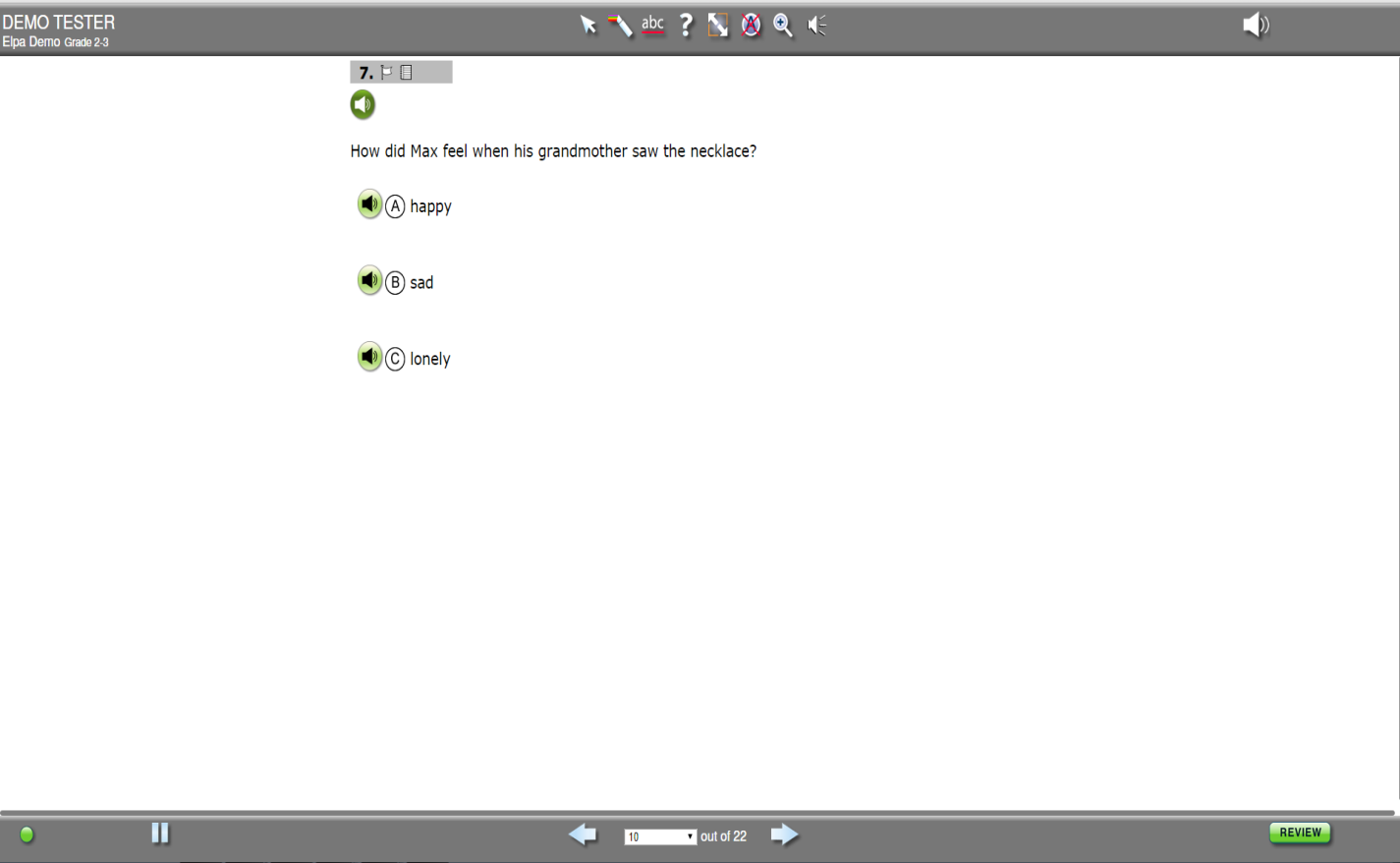




* The focal pictures of the question may need to be enlarged and/or centered.



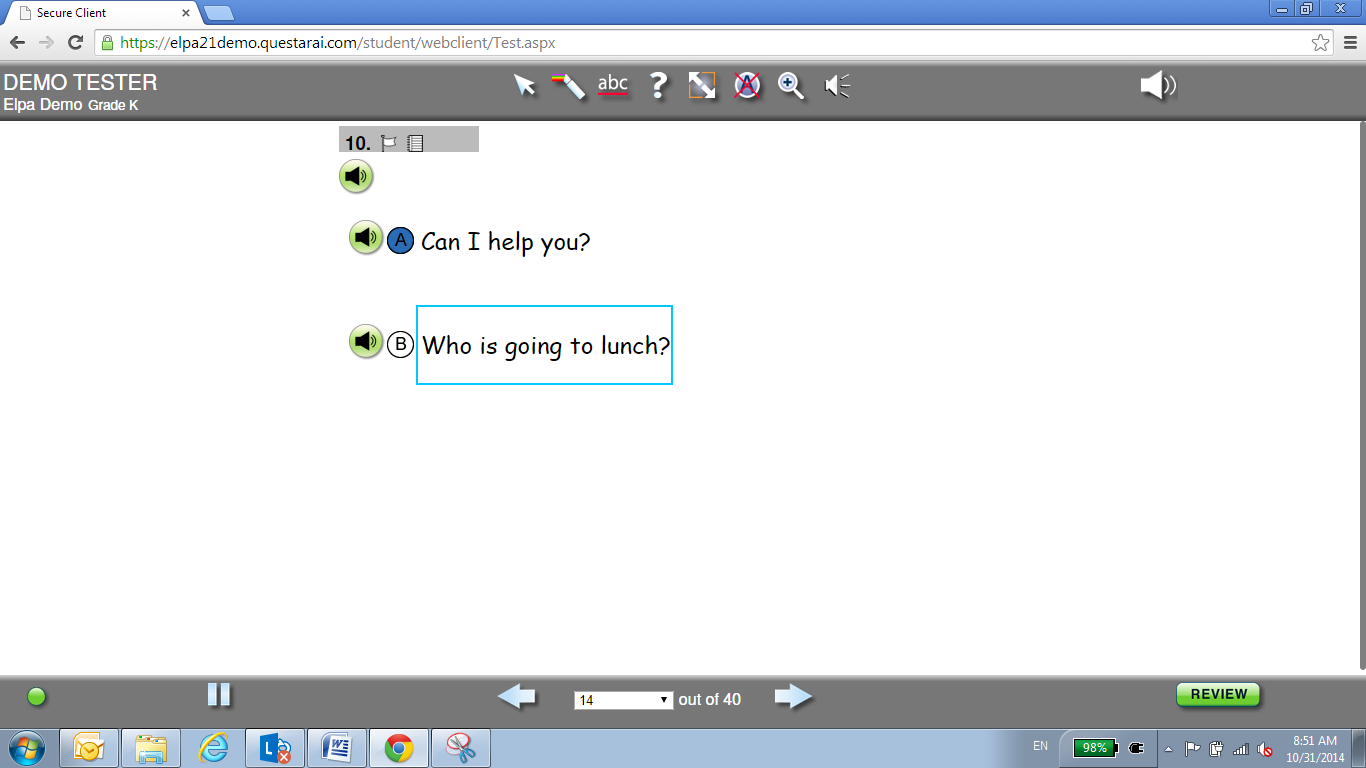
* The audio buttons sometimes caused confusion because they were next to the answer choices for all grade levels.



* Students tended to click and unclick on the answer choices repeatedly. They were unclear if they had finished answering after one click. Some students double clicked, and therefore unselected their answer or unintentionally skipped the item (when double clicking on the navigation arrows). Some kind of reminder may be necessary to let the student know that s/he did not respond to the item.



* Some students were confused by the box around the text-based multiple choice options because it appeared with a mouse over the choice.



* Students had difficulties using arrows at times to navigate between questions. Some students did not know to use the arrow to navigate. Some students clicked the arrow twice and missed items and needed to be told to go back to the previous page. Many students needed prompting or assistance to understand how to use the navigation arrows.

1. Technology-enhanced (TE) features in the items and technology features in the platform are referred to interchangeably in this report as these features are intertwined for test takers. In order to respond to the TE item types, test takers must perform specific interactions with the technology features, such as drag & drop, selecting from a drop-down menu, zone selection, and using recording functions. In responding to items, test takers use a range of technology features included in the platform (e.g., navigating test platform buttons, using a scroll bar) in addition to the features elicited by the items themselves (e.g., using a mouse, keyboarding, etc.) [↑](#footnote-ref-1)
2. The text regarding how recommendations and suggestions were and will be addressed was added by ELPA21 to the ETS report. [↑](#footnote-ref-2)
3. The text regarding how recommendations and suggestions were and will be addressed was added by ELPA21 to the ETS report. [↑](#footnote-ref-3)