WIDA™

Technical Report

No. TR-2019-2
March 2019

Investigating K-12 English Learners' Use of Universal Tools Embedded in Online Language Assessments

Ahyoung Alicia Kim, Ed.D., Meltem Yumsek, M.S., Mark Chapman, Ph.D., & H. Gary Cook, Ph.D.

Wisconsin Center for Education Research, University of Wisconsin–Madison

Wisconsin Center for Education Research
SCHOOL OF EDUCATION
UNIVERSITY OF WISCONSIN–MADISON

Contents

List of Tables

List of Figures

Executive Summary

English learners (ELs) with disabilities face challenges to meaningful participation in educational institutions. Today, more than 10% of the Kindergarten to Grade 12 (K-12) ELs in the U.S. are identified as having one or more disabilities (U.S. Department of Education, 2014). The Every Student Succeeds Act entitles ELs with disabilities to one or more accommodations when taking English language proficiency (ELP) assessments. Test designs should allow ELs with disabilities to meaningfully participate in the assessments, so they may demonstrate their ELP in a manner consistent with their peers. Furthermore, assessments should employ features of universal design, such as a text highlighter and magnifier, that improve accessibility for all test-takers. These accessibility features, also known as universal tools, are designed to provide the necessary support for the general EL population, including ELs with disabilities (Willner & Monroe, 2016). It is important to understand how ELs use accessibility features, as both those with and those without disabilities are student populations with special needs. However, no known studies have investigated this topic in the ELP assessment context, suggesting the need for research.

This study examined how Grades 1-12 ELs with and without disabilities used online accessibility features during an ELP assessment. The test studied is ACCESS for ELLs (hereafter ACCESS) that is administered annually across 39 U.S. states and territories to measure the four language domains of listening, speaking, reading, and writing. Test scores are used for making high-stakes decisions about ELs, such as placement and reclassification. Approximately 1.3 million ELs' test telemetry data (i.e., records of test-takers' online interactions during the test) were analyzed, with 11% of data capturing the test-taking activity of ELs with disabilities. The study focus was on ELs' use of several accessibility features embedded in the online ACCESS platform: *Color Overlay, Color Contrast, Help Tools, Line Guide, Highlighter, Magnifier*, and

Sticky Notes. To explore the degree to which ELs use the accessibility features, descriptive and frequency analyses of the telemetry data were conducted for each feature across all ELs, and for ELs with and without disabilities. This study also examined the relationship between students' ELP and their use of accessibility features as well as individual items that triggered the increased use of one or more accessibility tools. To compare the use of accessibility tools between ELs with and ELs without disabilities, and across three proficiency levels (beginner, intermediate, and advanced), Mann-Whitney U tests were conducted due to non-normal distributions. Effect sizes were reported, in addition to the significance of group differences.

Findings show that ELs as a whole generally used the Line Guide, Highlighter, and Magnifier more frequently than other accessibility features, and they used the Help Tools the least. Use of accessibility features was more common in the selected-response listening and reading domains, which were administered prior to the constructed-response speaking and writing sections. The comparison of ELs with disabilities to those without disabilities revealed that a higher percentage of ELs with disabilities activated the accessibility features across all domains. Although the difference in the use of some features between the two groups was statistically significant, effect sizes were small. Findings regarding the three proficiency groups showed that intermediate and advanced ELs demonstrated a higher percentage of tool use across the four domains than beginners. Again, although there were significant differences in tool use across the proficiency levels, the effect sizes were small. In addition, results suggest that some item types triggered the increased use of accessibility tools, with many ELs using a combination of tools to respond to specific item types. Overall, study findings indicate the usefulness of the accessibility features that are embedded in online tests, suggesting that such features may provide the intended supports for special populations of students, including ELs with disabilities.

1. Introduction

English learners (ELs) in the Kindergarten to Grade 12 (K-12) context in the U.S. continue to be a rapidly growing population. Overall, ELs constituted 9.5% of the public school student population in 2014-2015, and in states like California and Nevada, the EL population reached 17-20% (National Center for Education Statistics, 2018). The National Education Association (n.d.) anticipates that by 2025, the overall percentage of ELs will rise to 25%. The EL population is also diverse and includes students from various ethnic and cultural backgrounds and age groups. One overlooked subpopulation of ELs is those identified with disabilities. The number of students identified with disabilities is increasing nationwide (Guzman-Orth, Laitusis, Thurlow, & Christensen, 2016), and in 2015 about 14.7% of the ELs in public elementary and secondary schools were reported to have a disability (National Center for Education Statistics, 2018).

ELs face a myriad of challenges in the classroom and assessment contexts, and often fall behind their native English speaker counterparts. It is a complex task to identify the cause of this gap due to the multiple factors that may contribute to this problem. This task becomes more difficult when considering ELs with disabilities, as language- and disability-specific factors are often convoluted (Zehler, Fleischman, Hopstock, Pendzick, & Stephenson, 2003). However, federal law, previously in the form of the No Child Left Behind (NCLB) Act and now under the Every Student Succeeds Act (ESSA), mandates maximizing the inclusion of not only ELs but also ELs identified with disabilities in assessments (U.S. Congress, 2015; U.S. Department of Education, 2014). Thus, it is the responsibility of test developers to ensure valid interpretations of test scores for these special populations. Reducing the impact of extraneous factors (i.e.,

disability-related factors) or construct-irrelevant variance increases the chance that these student groups can successfully show their true abilities in the areas the assessments intend to measure.

The *Standards for Educational and Psychological Testing* (hereafter, the *Standards*; American Educational Research Association [AERA], American Psychological Association [APA, & National Council on Measurement in Education [NCME], 2014) define construct-irrelevant variance as "the degree to which test scores are affected by processes that are extraneous to the test's intended purpose. The test scores may be systematically influenced to some extent by processes that are not part of the construct" (p. 12). Construct-irrelevant variance is attributed to various factors such as (1) content (e.g., overly linguistically complex tasks), (2) response formats (e.g., new question and response types), and (3) test context and environment (e.g., complicated instructions, computer skills). Construct-irrelevant variance not only presents a primary validity threat by contaminating the test scores and score interpretations (Messick, 1996), but also poses a fairness risk. According to the *Standards* (AERA et al., 2014), fairness means maximizing "the opportunity for test takers to demonstrate their standing on the construct(s) the test is intended to measure" (p. 51). A fair test does not include characteristics irrelevant to the test construct. One of the ways to minimize construct-irrelevant variance for ELs, including ELs with disabilities, is to provide adjustments to tests and testing situations under certain circumstances.

With the advancement of technology, online assessments and computer-based tests have incorporated numerous features and tools (e.g., virtual highlighters, dictionaries) that could enhance learners' test-taking experience and increase accessibility to tests. These accessibility features, also known as universal tools, are intended to provide the necessary support for all student populations, including ELs, so that all students may fully demonstrate their language

abilities (Willner & Monroe, 2016). By definition, universal tools are "selectable embedded features or hand-held instruments used to carry out a particular purpose. Universal tools may either be embedded in the online test or provided to [ELs] by test administrators for online or paper tests" (Willner & Monroe, 2016, p.3). Note that the concept of accessibility, as implemented through features like universal tools, differs from the concept of accommodations. The latter is a narrower term that refers to supports provided to only a subgroup of students, such as ELs with disabilities (refer to the Literature Review below for more detail).

The focus of this study is on the use of universal tools in English language proficiency (ELP) assessments by ELs, including those with disabilities. Because universal tools have been designed to serve particular functions and address students' needs, it is important to examine test-takers' use of the tools during test taking to know whether the tools are truly serving their intended purpose. Moreover, it is critical to investigate the accessibility of ELP assessments as ESSA requires ELP to be included as a success indicator in accountability models (Council of Chief State School Officers [CCSSO], 2016). Prior to ESSA, content-based assessments, such as math and English language arts assessments, were part of accountability but ELP assessments were not. Now, ELP exams are used not only for placement and classification decisions, but also for accountability purposes. It is thus of the utmost importance to explore whether ELP assessments are accessible and allow ELs to show their true abilities. This research around accessibility is fairly new, and baseline studies are virtually non-existent, particularly in the context of K-12 ELP assessments. Moreover, ELs with disabilities have often been overlooked in both practice and research. The present study aims to address these gaps.

2. Literature Review

2.1. Defining Accessibility and Accommodations

The notion of accessibility is an inclusive term that considers all target student populations, including ELs and students with disabilities. In order to define accessibility, it is necessary to discuss how it differs from accommodations. By definition, accommodations refer to embedded or conditional adjustments to a test or testing situation (i.e., the test presentation, format, environment, or administration) without altering the test construct (AERA et al., 2014; Pennock-Roman & Rivera, 2011). Accommodations are applied to increase equity for student populations with specific needs (Abedi, 2017) and improve their access to an assessment (Pennock-Roman & Rivera, 2011). While accessibility is for all students, accommodations are specifically designated for special populations such as students with disabilities. In contrast to current usage, the term accommodations has traditionally been used to refer to supports that are offered to ELs, particularly in the context of content-based assessments.

For the EL population taking content assessments, accommodations are intended to address their second-language needs (Abedi, 2017) and provide a fair opportunity for them to demonstrate their knowledge and ability (Abedi, Hofstetter, & Lord 2004). Accommodations are categorized as either direct or indirect linguistic supports (Rivera, Collum, Shafer Willner, & Sia, 2006; Shafer Willner, Rivera, & Acosta, 2008). Direct linguistic supports are associated with alterations to the test language or content. Examples include translation of test items into an EL's first language, simplified word or syntax choices, and allowed use of dictionaries or glossaries (Cawthon, 2010). Indirect linguistic supports involve changes to testing or administration conditions (e.g., extended time) designed to enhance ELs' language processing without changing the test itself (Crotts-Roohr & Sireci, 2017). Direct linguistic supports might be helpful for assessments that measure constructs other than language. However, there is concern over the

validity of score interpretations of ELP tests if linguistic supports are provided to test-takers working through items that actually test students' vocabulary skills or language as a construct (Abedi, Leon, & Kao, 2008). In addition, the use of accommodations for ELs has been criticized for not being responsive enough. Several researchers (Acosta, Rivera, & Shafer Willner, 2008; Abedi, 2017) put forth that the accommodation concept and existing strategies were borrowed from practices used with students with disabilities and are not ideally suited to meet ELs' needs (Rivera et al., 2006; Shafer Willner et al., 2008).

Moreover, in recent years, scholars have challenged the use of accomodations as the principal strategy to meet students' needs (Shafer Willner & Rivera, 2011). There has been a shift in the discussion from accommodations to accessibility, as reflected in the discussion in the *Standards* (AERA et al., 2014). Accessibility is different from accommodations in the sense that it embraces a wider population by allowing for a wider range of students to demonstrate their true knowledge or abilities without being hindered by construct-irrelevant features. Accessibility derives from the universal design framework, which offers an approach to test development and assessments that maximizes accessibility and validity from the onset of development, with consideration for all intended students, regardless of age, socio-economic status, disability, or linguistic or cultural background (Hansen & Mislevy, 2006; AERA et al., 2014).

Universal design and accessibility are concerned with numerous methods of representation; action and expression; and engagement in the design stages (Thurlow, Lazarus, Albus, & Hodgson, 2010). The principles include recognition of an inclusive target population, a precise construct definition, accessible and bias-free items and tasks, clear and straightforward instructions, and maximum readability, comprehensibility, and legibility (i.e., font, size, style, spacing, color, contrast) (Thompson, Johnstone, & Thurlow, 2002). In this regard, the

accessibility features embedded in an assessment include, but are not limited to, color and contrast modification, text-to-speech functions (i.e., the text is read aloud and visually highlighted), font enlargement, screen magnification, online dictionaries, spelling and grammar checkers, and extended time (Hansen & Mislevy, 2006). These features can assist students in meeting the requirements not directly related to the construct of the assessment. For instance, pop-up glossaries might address content accessibility challenges for ELs in a math assessment where vocabulary knowledge is deemed subsidiary.

Universal design presents numerous advantages. Considering accessibility in the early stages of test design can reduce construct-irrelevant variance and ensure more valid, reliable, and fair test results (Guzman-Orth et al., 2016). Any item or test feature or format that hampers measuring the construct or that introduces construct-irrelevant variance is discouraged within the universal design framework (AERA et al., 2014). Universal design also aims to overcome certain barriers that prevent student engagement with tasks. Prioritizing accessibility means offering test-takers multiple avenues of interacting with information (e.g., via manipulatives like a magnifier or color contrast tools), thereby diminishing difficulties in recognizing the information presented as text, graphics, or oral language and increasing access to content (Russell, Hoffman, & Higgins, 2009). For these reasons, assessments employing the universal design framework "hold great promise for accessibility barriers" for the EL population (Liu & Anderson, 2008, p.168).

2.2. Accessibility Features Embedded in Computer-Based Assessments

Computer-based testing has become popular partly due to a strong initiative from federal funding and legislation (i.e., Race to the Top, ESSA) that support innovative assessment design. Computer technology opens up the possibility of eliminating construct-irrelevant variance for all

students in the early design process (Dolan et al., 2010) and incorporating built-in accessibility features (Thurlow et al., 2010). Computer-based testing is powerful because test-takers can use multiple features simultaneously, which optimizes accessibility. Hansen and Mislevy (2006) advocate for accessibility features to be built into computer-based test platforms to reduce barriers specifically for ELs and students with disabilities, and to support those populations so they can fully demonstrate their abilities.

Accessibility features embedded in ELP assessments differ in nature from those in content assessments. The former tend to support the test-taking experience in general rather than providing explicit linguistic support, such as translating the prompt from English into students' home language, as ELP assessments are designed to measure language ability. Computer technology allows test developers to maintain and/or introduce universal tools (i.e., note-taking, outlining, or other "scratch paper" tools; highlighters; and line guides) in ELP assessments that are related to processing skills and meaning-making, and which may encourage ELs to continue using their existing study and learning skills; this characteristic of technology makes online assessments more accessible to ELs without impacting the focal construct. Additionally, computer-based testing enables test developers to track the extent to which accessibility features are used and obtain data to explore the cognitive processing (e.g., the response time) of test-takers (Crotts-Roohr & Sireci, 2017).

Investigating accessibility in the computer-based testing context is paramount for understanding the effectiveness and validity of the available accessibility features. Currently, there is limited evidence to support the effectiveness of accessibility tools in the computer-based testing environment. Therefore, researchers are accumulating such evidence (AERA et al., 2014). The following section provides an overview of recent relevant research.

2.3. Previous Research on Accessibility Features for English Learners

Previous studies on the topic of accessibility have focused on students with special needs (Russell, Johnstone, Higgins, and Hoffman, 2008; Russell, Hoffman, & Higgins, 2009). For instance, Russell et al. (2008) explored the effectiveness of accessibility features (i.e., read-aloud, magnification, color contrast) by comparing the performance of students with and without special needs in Grade 6 and Grade 9. Both student groups took two content-focused assessments, with or without accessibility features. The results showed that accessibility features had a positive impact on the performance of students with special needs while deteriorating the performance of students without any special needs. Nevertheless, most of the students in the study found the tools user-friendly, and 85% of the students were eager to make use of the tools if presented in future tests.

Although there has been a plethora of research on accommodations, studies on accessibility features for ELs are still emerging (AERA et al., 2014). The majority of the research on computer-based accessibility features pertains to content-focused assessments for ELs (Abedi, Bayley, Ewers, Mundhenk, Leon, & Kan,, 2012; Cohen, Tracy, & Cohen, 2017; Crotts-Roohr & Sireci, 2017; Carr, 2008; Kopriva, Winter, Triscari, Carr, Cameron, & Gabel, 2013; Kopriva, Triscari, & Carr, 2014).

Most studies (Abedi, 2009; Abedi et al., 2012; Cohen et al., 2017; Crotts-Roohr & Sireci, 2017; Kopriva et al., 2013; Kopriva et al., 2014) focused on pop-up glossaries or dictionaries as well as some emerging tools for font manipulation and paraphrasing. For example, Abedi (2009) compared the performance of ELs and non-ELs in Grade 4 and Grade 8 with and without the following supports: (1) a pop-up glossary, (2) a customized English dictionary, (3) extra testing time, and (4) small-group testing. The glossary and dictionary features were embedded into a content-focused mathematics assessment. (Items were compiled from the National Assessment

of Educational Progress and Trends in International Mathematics and Science Studies.) This study showed that computer-based testing with linguistic support features was the most effective approach to making the assessment accessible for ELs. Students with access to these features performed better on the assessment. Furthermore, these features did not have a differentiated impact on the non-EL group, which further supports the validity of this approach.

More recently, Cohen, Tracy, and Cohen (2017) investigated the effectiveness and validity of the pop-up glossary feature within a computer-based math and English language arts assessment given to Grade 3 and Grade 7 EL and non-EL students. The glossary benefited the EL population; however, results were somewhat inconsistent. Although Grade 7 ELs' performance on the English language arts assessment improved when the test-takers had access to glossaries, their math achievement was hindered substantially due to the glossary use. Similarly, Grade 3 students' performance was slightly negatively affected by the glossary tool, revealing that pop-up glossary tools may obstruct the performance of younger learners.

These studies reveal that pop-up glossaries and dictionaries are commonly used as support tools for ELs in content-focused assessments. Some additional supports include read-aloud features, paraphrase tools, color setting options, and magnification devices. While most studies found supports provided to ELs to be effective, there is also conflicting evidence. Especially for younger learners or for students who do not need them, these tools might have deteriorating effects.

In another line of research, studies have examined how accessibility features affect the performance of ELs of varying proficiency levels in the context of content-focused assessments. While some studies (Emick & Kopriva, 2007; Kopriva et al., 2013) found evidence of high proficiency ELs performing better when given accessibility features, others (Carr, 2008; Crotts-

Roohr & Sireci, 2017; Kopriva et al., 2014) suggest otherwise. For instance, Emick and Kopriva (2007) found evidence of the usefulness of universal tools (plain or simplified language and visuals) for high proficiency ELs. High proficiency ELs also outperformed native English speakers on a biology test (Kopriva et al., 2013). On the other hand, Crotts-Roohr and Sireci (2017) focused on ELs of two proficiency levels (i.e., mid and high proficiency) and non-EL students in high school math and history assessments. They investigated the extent to which students used a pop-up glossary and a paraphrasing tool and how those tools impacted both performance and response time. Although ELs made significantly more use of the two features in both assessments, their use decreased as they progressed through each test. Also, the mid-proficiency group used the features more than the high-proficiency and non-EL groups. While the response time of ELs extended in conjunction with the use of these two features in the history assessment, the same pattern was not observed in the math assessment. These studies show that accessibility features have been generally effective for ELs in the context of content-focused computer assessments. However, the relationship between different proficiency levels and the effectiveness of accessibility features was inconsistent, suggesting more research on the topic is needed.

Some researchers have attempted to investigate expert or educator perceptions of accessibility features. For instance, Liu and Anderson (2008) surveyed experts to understand their opinions about a list of accessible testing considerations specified in existing research literature. With respect to universal design considerations specifically for computer-based assessments, they found that sufficient contrast between background and text and a background that does not interfere with readability were among the features that received the highest

rankings from experts. However, participant views varied largely with respect to the importance of the use of manipulatives, assistive technologies, and dictionaries.

Compared to the literature on accessibility features in content-focused assessments, existing research on such supports in ELP assessments is limited. Furthermore, the majority of the prevailing research is concentrated on foreign language assessment in higher education rather than K-12 contexts (e.g., Choi & Cho, 2016; Frankenberg-Garcia, 2005, 2011; Oh, 2018). These studies typically focus on reading and writing assessments and linguistic tools to support test taking. For instance, Oh (2018) examined the nature of test-takers' use of the spelling, grammar, dictionary, and thesaurus tools in a writing assessment and evaluated the validity of these tools for ELP assessments. The study adopted a mixed-method design by analyzing scores from three writing tasks and video recordings of the test-taking process of adult second language speakers of English (n = 120). The performance of the participants at different proficiency levels with and without access to the universal tools was compared. The findings show that when test-takers had access to a universal tool, especially the spelling and reference tools, the writing test still had high reliability and dependability providing further support for the validity of these tools. Findings also demonstrated that inclusion of these tools increased interactivity and the authenticity of tasks.

Studies examining the use of universal tools in K-12 ELP assessments are virtually nonexistent. Kim, Monroe, and Lee (2018) investigated EL educators' perceptions of universal tools and how their students make use of them. In a mixed-methods study, the researchers surveyed K-12 educators from 30 states (n = 377) and conducted follow-up interviews with nine of those educators. The authors indicate that educator perceptions of universal tools vary, and educators valued certain tools more than others. The tools educators valued most were a

highlighter, a line guide, and an underlining tool. The educators also attributed the frequency with which students made use of accessibility features to a variety of factors, including familiarity with computers, length of residence in the U.S., grade level, and special needs.

Universal tools in the computer-based K-12 ELP assessment context might provide effective supports for ELs, as inferred from the few studies discussed above. However, more studies need to be conducted for researchers to learn the extent to which ELs themselves use these tools during ELP assessments. Moreover, very little is known about how ELs with disabilities use accessibility features during testing. When assessments are computer-based, there is a wider variety of item presentation formats and response functionalities, which creates more accessibility challenges for both ELs and ELs with disabilities (Guzman-Orth et al., 2016). Furthermore, it is unknown how the use of universal tools relates to language proficiency. Abedi (2014) asserts that proficiency level is among the most critical criteria in making appropriate accessibility and accommodation decisions for ELs.

In conclusion, accessibility research for ELs is still emerging. Although existing studies have investigated accessibility in content-based assessments, accessibility in ELP assessments is relatively underexamined specifally in the K-12 context and therefore requires further research.

2.4. Present Study

The main purpose of the current study is to examine the use of computer-based accessibility features among K-12 ELs, including ELs with disabilities (also referred to as ELs with an Individualized Education Program [IEP]) in an online ELP assessment. The study elucidates the extent to which students use the accessibility features embedded in a particular computer-based ELP assessment. Consequently, it examines the validity of accessibility features by exploring their use. The study also seeks to scrutinize (a) how use of the tools varies between

ELs in general and ELs with IEPs, (b) the relationship between students' ELP and their use of accessibility features, and (c) how item types affect the use of accessibility features. To this end, student data from the ACCESS for ELLs assessment is analyzed (details of ACCESS are provided in the Methods section) to answer the following research questions:

1. To what extent do ELs, including ELs with disabilities, use the accessibility features embedded in an ELP assessment?
2. To what extent do ELs with and without disabilities differ in their use of the accessibility features embedded in an ELP assessment?
3. To what extent do ELs at different proficiency levels vary in their use of the accessibility features embedded in an ELP assessment?
4. To what extent do different item types and features affect ELs' use of the accessibility features embedded in an ELP assessment?

3. Methodology

This section describes the study's context, data, and analysis procedures.

3.1. Context of the Study

In examining the use of accessibility features among ELs, including ELs with IEPs, this study focuses on ACCESS, a large-scale standardized ELP assessment that is widely used in the K-12 context. ACCESS is developed by WIDA in collaboration with the Center for Applied Linguistics, and it is annually administered to over 2 million K-12 ELs in 39 states and territories. ACCESS measures the academic ELP development of K-12 students in the four language domains of listening, reading, speaking, and writing.

ACCESS is offered in paper and online formats. The context of this study is the online format, which is available for five grade-level clusters: Grade 1, Grades 2-3, Grades 4-5, Grades

6-8, and Grades 9-12.[1] Except for the Grades 1-3 writing section, ACCESS Online is delivered completely via computer by default. However, states and school districts do have the option of allowing Grades 4-5 students to handwrite or keyboard their responses, and test administrators can provide the handwriting option to Grades 6-12 students due to special needs. Thus, the writing samples included in this study, which examines only responses to a computer-based format of the test, make up a smaller data set than that of the other three domains. The listening and reading domains of the ACCESS Online test are presented in a multi-stage, adaptive format, and all the domains have a tiered structure, reflecting the targeted difficulty of the test (i.e., Tiers A and B/C for listening, reading, and writing, with Tier A being the easiest; Pre-A and Tier A/B/C for speaking, with Pre-A being the easiest and C being the hardest). The number of items or tasks changes respective to each particular domain and tier.

ACCESS was developed under the principles of universal design to ensure the inclusion of all students (WIDA, 2018). A large number of accessibility and accommodation features are embedded in the ACCESS Online test platform to support students' access to the test content. This study focuses on eight of those accessibility features: Color Overlay, Color Contrast, Highlighter, Line Guide, Magnifier, Help[2] [Help (General) and Help (Tools)], and Sticky Notes. Although some of these tools are self-explanatory, a detailed description of each is provided in Table 1 below. Also, a screenshot of the ACCESS Online test is shown in Figure 1 to demonstrate how the universal tools are presented in the testing environment. The universal tools are available throughout each domain and can be accessed for all individual items, except for the Sticky Notes, which is exclusive to the writing domain. Students can activate the tools an

[1] Kindergarten students are given only the paper version of ACCESS due to the challenges of young learners completing a test via an electronic platform.
[2] The Help tool was analyzed as two separate options—Help (General) and Help (Tools)—but the conclusions are sometimes reported together as no significant differences were found.

unlimited number of times at any point during the test. Students can also use a combination of tools, depending on their needs, for each individual item. Online practice sessions are available to educators to help them familiarize students with the accessibility features.

Table 1. Description of universal tools examined in this study

ACCESS Online Universal Tools	Availability
 Color Overlay allow test-takers to manipulate the text color and the background color that appears behind the text, graphics, and response areas. There are 6 pre-defined combinations.	All 4 domains
 Color Contrast allows test-takers to manipulate the contrast between text and background by selecting a background color. There are 6 pre-defined background colors.	All 4 domains
 Highlighter allows test-takers to mark parts of the text presented.	All 4 domains
 Line Guide allows test-takers to drag a horizontal line across the lines of the text presented.	All 4 domains
 Magnifier allows test-takers to manipulate the graphic and text size, which can be enlarged to 1.5 or 2.0 times the default size.	All 4 domains
 Help gives test-takers more information about the universal tools, with 2 options: (1) “What’s This?” [which is referred to as Help (General) in the report] that describes how to use the Help tool, and (2) “Open Help” [which is referred to as Help (Tools)] that explains how to navigate the online test platform and activate the universal tools.	All 4 domains
 Sticky Notes gives test-takers a free-write space to organize ideas and plan their writing.	Writing

Figure 1. A sample test item in ACCESS Online showing universal tool buttons at the bottom of the screen

3.2. Data

The data for this study comes from the 2016-2017 administration of ACCESS Online. In addition to student test data, telemetry data was examined in this study. These data sets are described below.

3.2.1. Student Test Data

A total of 1,249,151 Grades 1-12 students from 37 states and territories took ACCESS Online during the 2016-2017 administration. About 43% of the students were enrolled in Grades 1-3. Overall, 11.7 % of the students who took the online form were reported to have a disability

based on their IEP status.[3] The ethnicity distribution of ACCESS Online test-takers shows that the majority of the tested EL population was Hispanic (64.2%). Among the students reporting their gender, 53.9% was male and 44.4% was female. The proficiency levels of the ACCESS assessment range from 1.0–6.0, and the average proficiency of test-takers in the 2016-2017 administration year was 3.4. The average proficiency for the aggregate EL group was highest in the listening domain (4.6), while the average lowest proficiency was in the speaking domain (2.9).

For the purposes of this study, ELs who were missing a proficiency level score in any domain were removed from the analysis. Their universal tool use data would be incomplete, as these students did not attempt all items. Thus, the sample sizes show variation across the four domains in the analysis. For the writing domain, Grades 1-3 students' data were excluded, as they did not complete the test in an electronic format and did not use any of the tools. For the same reason, data of the Grades 4-12 students who handwrote their responses due to district policy or individual needs were also removed from the analysis for the writing domain.

Additionally, for Research Question 2, which examines tool use among ELs and ELs with IEPs, six U.S. states and territories were removed from the data set (see Table 2). New Hampshire did not report the IEP status of any students. Data from the other states were removed because the percentage of ELs with IEPs in these states was remarkably small compared to the nationally reported average of 14.7% (NCES, 2018), suggesting that these states may not have fully reported ELs' IEP status.

[3] All public school students who have a disability and receive special education and other related services are required to have an IEP in order to improve their quality of education (U.S. Department of Education Office of Special Education and Rehabilitative Services, 2000).

Table 2. Percentages of ELs with IEPs in six states dropped from study

State	% of ELs with IEPs			
	Listening	Reading	Speaking	Writing
Delaware	1.2	1.2	1.1	1.1
Missouri	0.6	0.6	0.6	0.5
North Carolina	0.5	0.7	0.5	0.7
New Hampshire	0.0	0.0	0.0	0.0
Utah	0.1	0.1	0.1	0.1
Virgin Islands	0.6	0.6	0.6	0.7

3.2.2. Telemetry Data

Telemetry data is the information "collected behind the scenes via the web-based test engine regarding the actions test-takers execute while taking the test" (Data Recognition Corporation, n.d.). These data track the actions test-takers take in the computer-based environment, such as the activation of universal tools or a pausing of the test. Telemetry data also shows the duration of each activity, including the total time a test taker spends on each test item or screen. It presents rich and detailed information about students' processes while taking the test. The test engine starts recording the telemetry information once a test taker enters a module (i.e., one of the domains) and continues recording until the entire test is submitted and the test taker exits the system. All the actions are logged as event types in the telemetry data and a time for each event is captured in milliseconds. This study relies on telemetry information in order to understand students' interaction with the universal tools. Using the telemetry information, a data set—including variables that show the number of times each universal tool is activated by students in each domain—was created and analyzed.

3.3. Procedures for Data Analysis

In order to explore the extent which ELs use the universal tools built into ACCESS (Research Question 1), descriptive and frequency analyses were conducted on ELs' universal tool use. The analyses were carried out for each universal tool in each of the four language

domains; the main focus was on examining the percentages of students who used a given tool at least one time while completing the domain. The same analyses were undertaken for each of the five grade-level clusters to uncover the extent of universal tool use by students in different grades.

For Research Question 2, regarding the variation of universal tool use between ELs and ELs with IEPs, descriptive and frequency analyses were conducted separately on the two groups. Significance testing of differences between the two groups was conducted for each accessibility feature and the effect size was calculated. Due to non-normal distribution of tool use across the four domains and unequal samples sizes for each group, the Mann Whitney U test[4] was for significance testing found that the tool by ELs with and without IEPs among students was positively skewed. In addition to investigating data on ELs with IEPs, universal tool use by Plan 504[5] students and by students with different types of disabilities is explored.

With respect to Research Question 3, investigating the differences in universal tool use across proficiency levels, all students were first categorized into three groups based on their domain-specific levels: *Beginner* (Proficiency Levels 1-2), *Intermediate* (Proficiency Levels 3-4) and *Advanced* (Proficiency Levels 5-6). Frequency and descriptive analyses were then conducted for each of the three proficiency level groups in each domain. Significance testing of group differences was carried out and effect size was computed for each tool. Due to large

[4] The Mann Whitney U test is commonly known as the nonparametric analogue of the independent samples *t* test, which compares two groups (Howell, 2013). According to Howell, *t* tests and ANOVA are robust tests, and violations of assumptions (e.g., non-normality) have only a minor impact on the results. Still, the violation of assumptions mixed with unequal sample sizes for each group studied is a serious issue, as the test is less robust in these circumstances.

[5] Plan 504 is a federal civil rights law protecting all individuals with a disability, including students. The definition of disability in the 504 context is broad in comparison to the definition used for writing an IEP. For instance, any of the following can be considered a disability in Plan 504: physical or mental impairment, cosmetic disfigurement, anatomical loss, mental or psychological disorder, emotional or mental illness, and specific learning disabilities. Plan 504 students do not necessarily receive special education (U.S. Department of Education, 2016).

sample sizes, a non-parametric test was run to compare group differences; because there were three proficiency levels to compare, the Kruskal Wallis[6] test was used. Post-hoc pairwise comparisons were conducted using the Mann Whitney U test and effect size was computed for each pair. Additionally, frequency and descriptive analyses were conducted within each of the five grades clusters for each of the proficiency level groups to detect any differences among proficiency levels within a given grade.

While Research Questions 1-3 deal with the aggregate use of a given tool in each domain, exploring the use of tools for individual items might reveal important relationships as well. For instance, the use of some tools might depend on the type of item. In addressing Research Question 4, concerning the impact of item features on tool use, item-level tool use was analyzed for each grade-level cluster. Frequency and descriptive analyses were conducted, within each grade-level cluster, on the activation and use of each tool for each item. In addition, the average number of activation of a tool by students who used the tool at least once in the domain was calculated. The focus of this analysis was specifically on the percentage of students activating a given tool for a given item in each domain. Items (up to five items per domain) with a high rate of tool activation were flagged for more detailed exploration.

[6] The Kruskal Wallis test is a non-parametric equivalent of a one-way ANOVA test.

4. Findings

The findings are reported for each of the four research questions.

4.1. Findings from Research Question 1

To answer the first research question (To what extent do ELs, including ELs with disabilities, use the accessibility features embedded in an ELP assessment?), use of tools among all ELs and among all ELs by grade-level cluster is examined.

4.1.1. Use of Tools Among All ELs

Table 3 summarizes the universal tool use among all ELs, including ELs with disabilities, across the four domains of the assessment. The first column, "use," indicates the percentage of ELs that *activated* or *accessed* a particular tool at least one time (the report uses the term *use*, *activated*, and *accessed* interchangeably). Frequency of the tool use is represented in Figure 2. Descriptive data in Table 3 show the range and average number of tool activations or use for the group, as well as the, median, and variation in number of tool activations among the tool users only (not the entire sample set of test-takers). Frequency and descriptive statistics for the Sticky Notes are not applicable except for the writing section, as the Sticky Notes tool is available in the writing domain only. Following is a detailed description of the tool use in each domain.

In the listening domain, ELs used the Magnifier, Line Guide, and Highlighter features more frequently than Color Overlay, Color Contrast, and the two Help tools. In total, 9.3% of the ELs activated the Magnifier, 8.7% activated the Line Guide, and 4.7% activated the Highlighter at least one time during the whole listening section. Additionally, individual ELs accessed the Highlighter and Magnifier most frequently as indicated by the mean and median. Students activated the Magnifier 21 times on average, yet the median was two. Similarly, the average Highlighter activation was 12 times with a median of six times. It is clear that the activation of

all the tools was fairly positively skewed across domains, represented by the difference between the average and median values, particularly for the Highlighter and Magnifier. (Due to the skew of the data, median values are more closely examined in the findings.) Skewed data also suggests that some students may have consecutively clicked on the Highlighter or Magnifier button without making meaningful use of the tool. Comparatively, a low percentage of ELs activated the two color and two Help tools. The Help (tools) tool had the lowest access rate (1.8%). A slightly higher percentage of students activated the Help (general) tool (3%). A similar percentage of students accessed the Color Overlay (2.5%) and the Color Contrast (2.3%) tools. However, a majority of the activations of the Color and Help tools was a singular occurrence as indicated by the median.

For the reading domain, the Highlighter was the most accessed universal tool. The percentage of users increased to 11.1% (*Med*. = 9) from 4.7% in the listening domain. The percentage of ELs using the Magnifier remained the same as listening at 9.3% (*Med*. = 2). The Line Guide was expected to attract more students in the reading domain than in listening (due to the inclusion of reading texts), yet the use of the tool dropped slightly to 7.7%. However, the median of the use was higher (*Med*. = 2). The two Help tools were the least accessed features among ELs. In comparison to the listening domain, the activation of the Help tools in the reading domain decreased notably to 0.8% for Help (general) and 0.5% for Help (tools).

There was a general drop in the frequency of tool activation in the speaking and writing domains compared to the listening and reading domains. In the speaking domain, the tool activation was higher for the Magnifier (5%; *Med.* = 2) and the Line Guide (4.5%; *Med.* = 1) than for the other tools. Highlighter access dropped considerably to 3.8% (*Med.* = 5). The activation of the two Color tools also dropped to 1.5%. The two Help tools were again the least

accessed features. Overall, about 1% of the students activated these tools and almost all students were one-time users.

In the writing domain, the Highlighter was the most activated tool (5.3%; *Med*. = 6), followed by the Magnifier (4.9%, *Med.* = 2). Also, 4% of ELs used the Line Guide and Sticky Notes (*Med.* = 2); the use of the Sticky Notes tool was lower than expected despite its usefulness in providing students a place to organize their ideas before responding to the writing prompt. Again, the two Help features were the least frequently accessed tools. Less than 1% of the ELs activated them in the writing domain, which was similar to the activation rate in the speaking domain.

In conclusion, a higher percentage of students activated the Line Guide, Highlighter, and Magnifier than the other tools. The students maintained high use of the Magnifier across all domains. The Highlighter was the most accessed feature in both the reading and writing sections. Sticky Notes did not appeal to many students, contrary to our expectations. The use of some tools, like Color Overlay, Color Contrast, and the Line Guide, became less appealing to students as they progressed on the assessment from the listening domain to either the speaking or writing domain. The consistent decrease in the use of the Help tools could be partially attributed to the students' increased familiarity with the computer environment.

Table 3. Frequency and descriptive statistics for universal tool use across all domains among all ELs (including ELs with IEPs)

	Listening (n = 1,261,547)					Reading (n = 1,260,394)					Speaking (n = 1,237,802)					Writing (n = 657,934)				
Universal tools	Use %	Ran.	$\bar{X}$	Med.	sd	Use%	Ran.	$\bar{X}$	Med.	sd	Use %	Ran.	$\bar{X}$	Med.	sd	Use %	Ran.	$\bar{X}$	Med.	sd
Color Overlay	2.5	1-19	1.68	1.0	1.113	2.0	1-19	1.64	1.0	1.101	1.5	1-21	1.51	1.0	.959	1.7	1-19	1.52	1.0	.972
Color Contrast	2.3	1-19	1.79	1.0	1.242	2.1	1-29	1.78	1.0	1.267	1.5	1-36	1.58	1.0	1.094	1.9	1-21	1.66	1.0	1.138
Help (General)	2.9	1-22	1.12	1.0	.453	0.8	1-13	1.13	1.0	.475	1.0	1-10	1.10	1.0	.398	0.7	1-9	1.10	1.0	.387
Help (Tools)	1.8	1-10	1.11	1.0	.436	0.5	1-11	1.18	1.0	.563	0.8	1-11	1.15	1.0	.526	0.5	1-13	1.15	1.0	.535
Line Guide	8.7	1-870	1.93	1.0	3.256	7.7	1-311	2.90	2.0	3.65	4.5	1-192	1.89	1.0	2.041	4.0	1-113	1.63	1.0	1.52
Highlighter	4.7	1-968	12.38	6.0	23.477	11.1	1-973	17.41	9.0	28.842	3.8	1-1150	11.77	5.0	26.505	5.3	1-1191	12.21	6.0	24.942
Magnifier	9.3	1-240	20.65	2.0	38.079	9.3	1-295	23.8	2.0	42.809	5.2	1-236	45.67	2.0	63.744	4.9	1-464	16.32	2.0	26.266
Sticky Notes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4.1	1-117	2.02	1.0	2.313

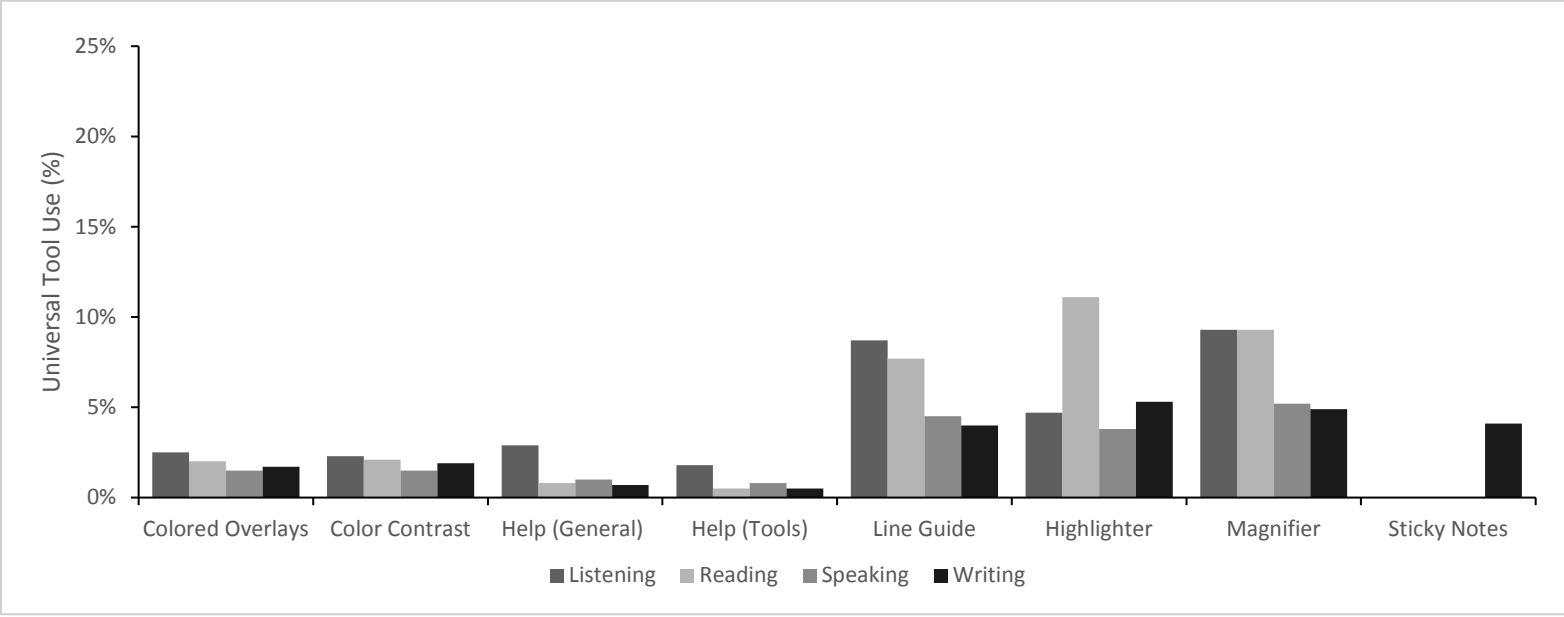

Figure 2. Frequency of universal tool use among all ELs in each of the four domains

4.1.2. Use of Tools among All ELs by Grade-Level Clusters

Given the anticipation that the universal tool use might change across grade-level clusters (e.g., young learners might make less use of the tools due to their limited computer knowledge), frequency and descriptive analyses were undertaken by grade-level cluster. The summaries of the findings for each domain are presented in Tables 4-7. Additionally, in order to visually compare the use of each tool in each grade-level cluster, bar graphs are provided in Figures 3-6. It must be noted that the sample sizes for each grade-level cluster show variation. While the Grade 1 cluster is the smallest, the Grades 2-3 students made up the largest data set, except in the writing domain, where Grades 9-12 was the largest and the Grades 4-5 the smallest data set.

The overall trend in the listening domain (Table 4 and Figure 3) shows that the Grades 6-8 students had the highest activation rate of universal tools among all grade-level clusters. They accessed the Magnifier (13.1%), Line Guide (12.5%), and Highlighter (7%) more than the other tools. Grades 4-5 students displayed the second highest access rate of the Magnifier (12.1%) and Highlighter (5.7%). Meanwhile, Grades 9-12 students had the second highest access rate of the Line Guide (11.1%). Despite variation in the percentages of students activating tools, the median for the Magnifier (*Med.* = 2) was the same for all grades except Grade 1 (*Med.* = 3). Similarly, the median number of activations of the two Color tools, two Help tools, and Line Guide were the same across all grades (*Med.* = 1). In general, Grade 1 ELs activated the accessibility features the least.

Table 4. Frequency and descriptive statistics for universal tool use in the listening domain for each grade-level cluster

	Color Overlay					Color Contrast					Help (General)					Help (Tools)				
	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*
Grade 1 (*n* = 175,826)	0.7	1-16	1.78	1.00	1.303	0.7	1-15	1.95	1.00	1.532	2.0	1-11	1.14	1.00	0.482	0.8	1-7	1.12	1.00	0.458
Grades 2-3 (*n* = 365,116)	1.5	1-12	1.76	1.00	1.183	1.5	1-19	1.84	1.00	1.296	2.6	1-8	1.11	1.00	0.391	1.3	1-9	1.09	1.00	0.399
Grades 4-5 (*n* = 216,937)	2.6	1-17	1.73	1.00	1.164	2.6	1-17	1.83	1.00	1.271	2.9	1-9	1.10	1.00	0.393	1.8	1-10	1.12	1.00	0.472
Grades 6-8 (*n* = 242,361)	5.0	1-19	1.71	1.00	1.128	4.6	1-16	1.84	1.00	1.271	3.8	1-11	1.13	1.00	0.460	3.0	1-9	1.12	1.00	0.474
Grades 9-12 (*n* = 261,307)	2.7	1-16	1.52	1.00	0.922	2.1	1-13	1.55	1.00	0.968	2.9	1-22	1.12	1.00	0.540	2.2	1-10	1.08	1.00	0.378

	Line Guide					Highlighter					Magnifier				
	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*
Grade 1 (*n* = 175,826)	4.4	1-111	1.72	1.00	2.234	2.3	1-579	12.55	6.00	23.605	5.6	1-240	41.68	3.00	47.165
Grades 2-3 (*n* = 365,116)	6.3	1-389	1.92	1.00	1.872	4.4	1-389	10.33	5.00	17.139	8.6	1-219	25.85	2.00	41.500
Grades 4-5 (*n* = 216,937)	9.0	1-66	1.96	1.00	1.791	5.7	1-462	10.33	5.00	17.189	12.1	1-188	18.14	2.00	36.089
Grades 6-8 (*n* = 242,361)	12.5	1-98	2.02	1.00	2.137	7.0	1-968	15.21	7.00	30.218	13.1	1-239	13.35	2.00	31.329
Grades 9-12 (*n* = 261,307)	11.1	1-870	1.89	1.00	5.387	3.5	1-657	13.42	6.00	25.352	6.8	1-168	16.56	2.00	34.385

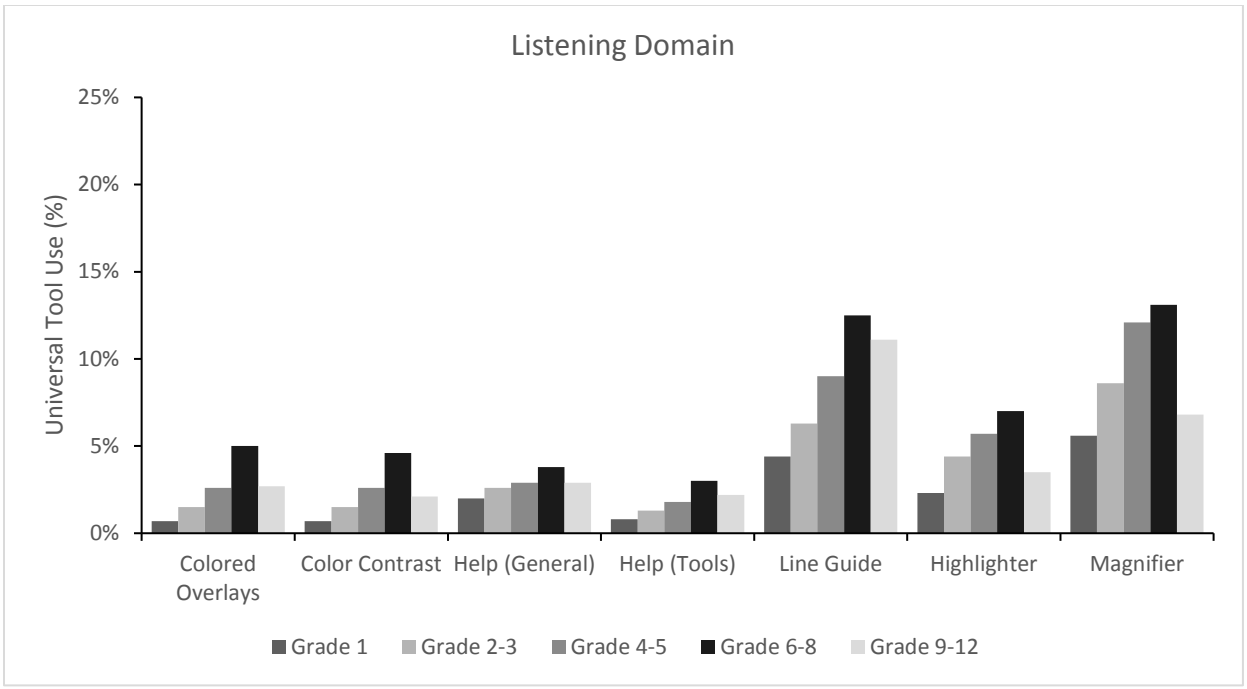

Figure 3. Frequency of universal tool use across grade-level clusters in the listening domain

In the reading domain (Figure 4 and Table 5), as compared to the listening domain, activation of the Highlighter increased, but activations of the Color tools, Line Guide, and Magnifier, as well as the two Help tools, generally decreased across all grade-level clusters. (The decrease in use of tools was minimal for Grade 1 students, but it was more noticeable for other grade-level clusters). One exception was the Grades 4-5 students, who showed higher percentages of access of the two Color tools, the Line Guide, and the Magnifier in the reading section as compared to the listening section. Also, the Grades 4-5 students showed the highest overall activation rate of the universal tools, including the Highlighter (17%), Magnifier (16.3%), and Line Guide (13.4%). In particular, use of the Highlighter tripled among Grades 4-5 students in the reading domain as compared to the listening domain (*Med.* = 9). Grades 6-8 students had the second highest percentage of tool use, including the Highlighter (13.8%), Magnifier (10.4%), and Line Guide (8.6%). The median number of times students accessed these tools was the same among the Grades 6-8 and Grades 4-5 populations. Grades 4-5 and Grades 6-8 students showed a

low percentage activation rate of the two Color and Help tools. Similar to the listening domain, the Grade 1 students displayed the lowest activation of the tools in the reading domain. Despite the low overall percentage of first grades' activation of the tool, their median use of the Magnifier was higher (*Med.* = 4) than that of students in the other grade-level clusters (*Med.* = 2).

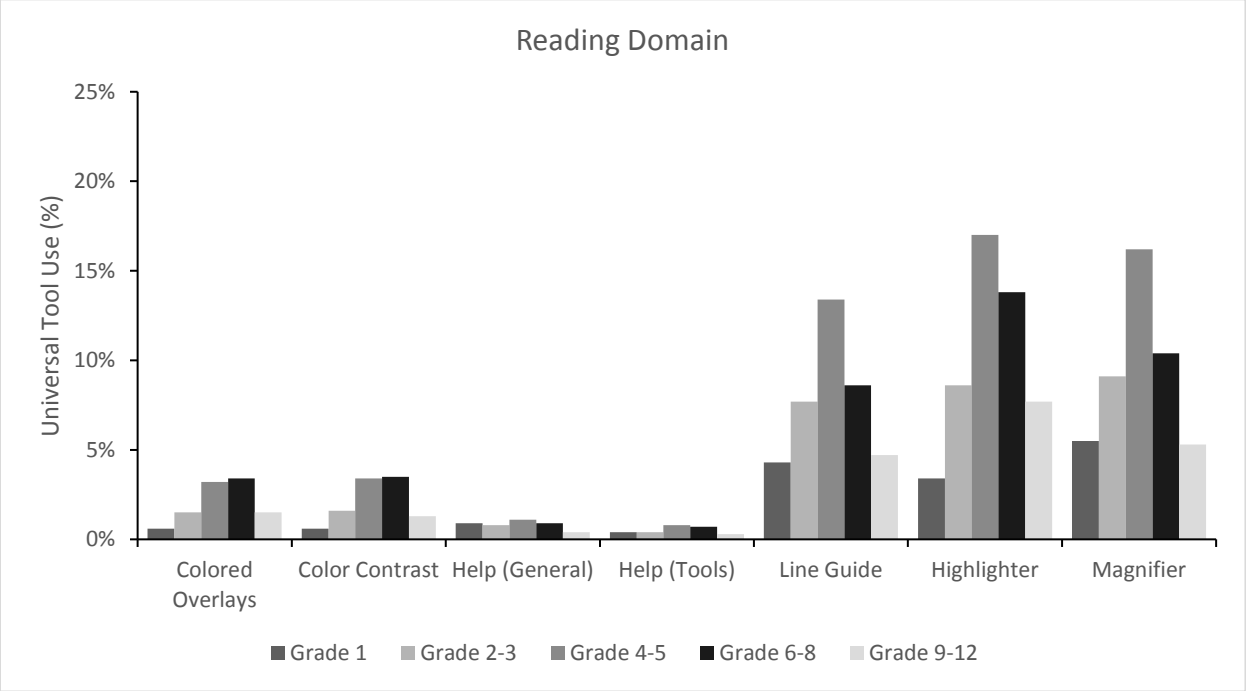

Figure 4. Frequency of universal tool use across grade-level clusters in the reading domain

Table 5. Frequency and descriptive statistics for universal tool use in the reading domain for each grade-level cluster

	Color Overlay					Color Contrast					Help (General)					Help (Tools)				
	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*
Grade 1 (*n* = 175,763)	0.6	1-11	1.62	1.00	1.109	0.6	1-10	1.68	1.00	1.091	0.9	1-8	1.15	1.00	.505	0.4	1-3	1.10	1.00	.342
Grades 2-3 (*n* = 364,939)	1.5	1-19	1.68	1.00	1.192	1.6	1-24	1.82	1.00	1.324	0.8	1-13	1.14	1.00	.515	0.4	1-8	1.18	1.00	.570
Grades 4-5 (*n* = 216,831)	3.2	1-13	1.74	1.00	1.192	3.4	1-16	1.91	1.00	1.383	1.1	1-5	1.14	1.00	.460	0.8	1-7	1.22	1.00	.618
Grades 6-8 (*n* = 242,241)	3.4	1-14	1.60	1.00	1.025	3.5	1-29	1.75	1.00	1.223	0.9	1-9	1.11	1.00	.451	0.7	1-11	1.20	1.00	.607
Grades 9-12 (*n* = 260,620)	1.5	1-15	1.49	1.00	.911	1.3	1-12	1.58	1.00	1.008	0.4	1-6	1.11	1.00	.378	0.3	1-6	1.12	1.00	.455

	Line Guide					Highlighter					Magnifier				
	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*
Grade 1 (*n* = 175,763)	4.3	1-66	2.24	1.00	3.211	3.4	1-568	13.83	6.00	22.891	5.5	1-168	48.04	4.00	53.489
Grades 2-3 (*n* = 364,939)	7.7	1-59	3.34	2.00	4.040	8.6	1-692	14.96	8.00	24.125	9.1	1-198	28.11	2.00	45.912
Grades 4-5 (*n* = 216,831)	13.4	1-90	3.36	2.00	3.825	17.0	1-704	17.06	9.00	26.787	16.2	1-222	16.93	2.00	36.405
Grades 6-8 (*n* = 242,241)	8.6	1-311	2.36	2.00	3.359	13.8	1-973	18.83	9.00	31.410	10.4	1-295	18.23	2.00	38.075
Grades 9-12 (*n* = 260,620)	4.7	1-93	2.10	1.00	2.525	7.7	1-761	20.58	9.00	35.229	5.3	1-160	24.08	2.00	42.5

In the speaking domain (Table 6 and Figure 5), the percentage of students activating each tool dropped in all grade-level clusters compared to the reading domain, with the exception of the Help tools by the Grades 9-12 students, which increased marginally. The decrease was most noticeable in activations of the Line Guide, Highlighter, and Magnifier tools. Grades 4-5 and Grades 6-8 students showed the highest activation of universal tools in the speaking domain. Access rates of the Color Overlay, Color Contrast, Help, Line Guide, and Highlighter tools was slightly higher among the Grades 6-8 students than the Grades 4-5 students. However, Magnifier activation was higher among the Grades 4-5 students than it was among the Grades 6-8 students. The median rates for accessing the two Color tools (*Med*. = 1), two Help tools (*Med*. = 1), Line Guide (*Med*. = 1), and Magnifier (*Med*. = 2. except for Grade 1) were the same across all grade-level clusters, despite variation in the percentages of students who accessed all the tools. On the other hand, the median access rate of the Highlighter varied across grades. Grade 1 students displayed lower activation of universal tools than other students in the speaking domain, just as they did in the listening and reading domains. Additionally, the median activation rate of the Magnifier was quite large for Grade 1 students (*Med.* = 108).

Table 6. Frequency and descriptive statistics for universal tool use in the speaking domain for each grade-level cluster

	Color Overlay					Color Contrast					Help (General)					Help (Tools)				
	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	sd
Grade 1 (*n* = 172,751)	0.5	1-14	1.58	1.00	1.010	0.5	1-13	1.70	1.00	1.211	0.7	1-10	1.10	1.00	.437	0.4	1-6	1.11	1.00	.467
Grades 2-3 (*n* = 361,122)	1.0	1-19	1.54	1.00	1.064	1.0	1-36	1.63	1.00	1.249	0.6	1-10	1.11	1.00	.450	0.4	1-11	1.15	1.00	.554
Grades 4-5 (*n* = 214,332)	1.9	1-21	1.53	1.00	.994	1.9	1-19	1.62	1.00	1.135	1.0	1-6	1.11	1.00	.394	0.8	1-7	1.18	1.00	.572
Grades 6-8 (*n* = 237,984)	2.7	1-13	1.51	1.00	.925	2.7	1-16	1.55	1.00	.963	1.7	1-6	1.10	1.00	.368	1.4	1-10	1.16	1.00	.546
Grades 9-12 (*n* = 251,613)	1.2	1-8	1.43	1.00	.825	1.0	1-17	1.47	1.00	1.029	0.9	1-6	1.09	1.00	.377	0.7	1-6	1.12	1.00	.434

	Line Guide					Highlighter					Magnifier				
	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*
Grade 1 (*n* = 172,751)	2.9	1-90	1.65	1.00	2.129	2.0	1-251	10.62	5.00	16.030	4.0	1-210	77.92	108.00	72.410
Grades 2-3 (*n* = 361,122)	3.8	1-73	1.90	1.00	1.892	3.5	1-432	8.95	5.00	15.348	5.1	1-218	53.40	2.00	66.923
Grades 4-5 (*n* = 214,332)	5.8	1-192	2.02	1.00	2.596	5.2	1-669	9.34	4.00	18.511	7.1	1-184	33.37	2.00	54.022
Grades 6-8 (*n* = 237,984)	6.2	1-44	1.95	1.00	1.873	5.6	1-1150	14.08	6.00	31.056	6.4	1-236	36.04	2.00	61.564
Grades 9-12 (*n* = 251,613)	3.9	1-39	1.75	1.00	1.569	2.7	1-951	17.16	6.00	42.799	3.3	1-208	41.70	2.00	57.391

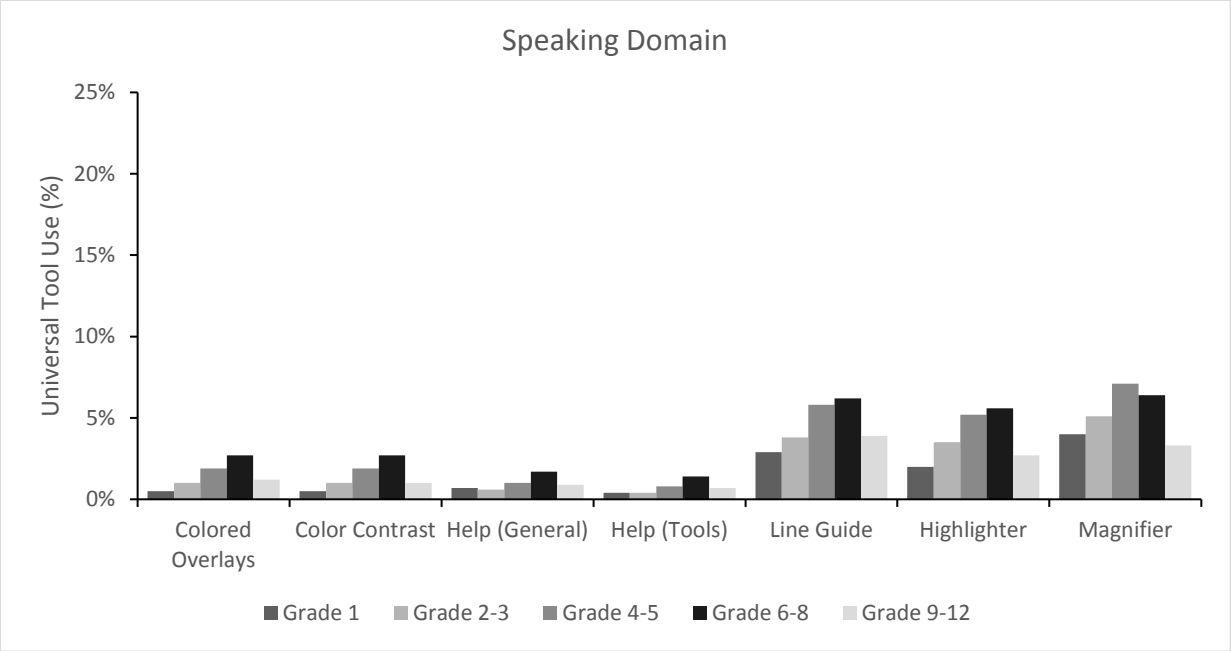

Figure 5. Frequency of universal tool use across grade-level clusters in the speaking domain

The analysis of the universal tools for the writing domain (Table 7 and Figure 6) consists of only three grade-level clusters because Grades 1-3 students handwrite their responses and the electronic universal tools are not available to them. As in the speaking domain, activation of each tool in the writing domain decreased in general in all grades compared to listening and reading sections. In the writing domain, the activation rate of all the tools was comparatively high among Grades 4-5 and Grades 6-8 students. A slightly higher percentage of Grades 6-8 students activated the Line Guide, Highlighter, Magnifier, and Sticky Notes tools than Grades 4-5 students, whereas they used the Color Overlay, Color Contrast, and Help tools similarly within these grade-level clusters. The tool activation was the lowest within the Grades 9-12 cluster. Specifically, the activation rate of two Color and two Help tools was less than 1%. Despite variation in the tool activation across the three grade-level clusters, the median activation of all tools but the Highlighter was the same across all grade-level clusters (*Med.* = 1)

Table 7. Frequency and descriptive statistics for universal tool use in the writing domain for each grade-level cluster

	Color Overlay					Color Contrast					Help (General)					Help (Tools)				
	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*
Grade 4-5 (*n* = 161,958)	2.1	1-17	1.64	1.00	1.158	2.4	1-18	1.81	1.00	1.315	0.9	1-5	1.10	1.00	.355	0.7	1-5	1.17	1.00	.531
Grade 6-8 (*n* = 239,611)	2.4	1-19	1.50	1.00	.923	2.7	1-21	1.64	1.00	1.085	0.7	1-5	1.10	1.00	.368	0.6	1-13	1.17	1.00	.614
Grade 9-12 (*n* = 256,373)	0.9	1-11	1.38	1.00	.755	0.8	1-11	1.47	1.00	.886	0.5	1-9	1.10	1.00	.443	0.4	1-4	1.09	1.00	.381

	Line Guide					Highlighter					Magnifier					Sticky Notes				
	Use %	Ran.	$\bar{X}$	Med.	sd	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*	Use %	Ran.	$\bar{X}$	Med.	*sd*
Grade 4-5 (*n* = 161,958)	4.8	1-53	1.65	1.00	1.382	7.0	1-962	10.69	6.00	19.224	7.1	1-350	15.95	2.00	28.061	5.6	1-60	1.10	1.00	.355
Grade 6-8 (*n* = 239,611)	4.1	1-113	1.65	1.00	1.809	5.9	1-677	13.47	7.00	25.345	5.3	1-219	16.04	2.00	24.925	4.8	1-117	1.10	1.00	.368
Grade 9-12 (*n* = 256,373)	3.3	1-41	1.59	1.00	1.248	3.6	1-1,191	12.16	5.00	29.996	3.2	1-464	17.29	2.00	25.658	2.4	1-61	1.10	1.00	.443

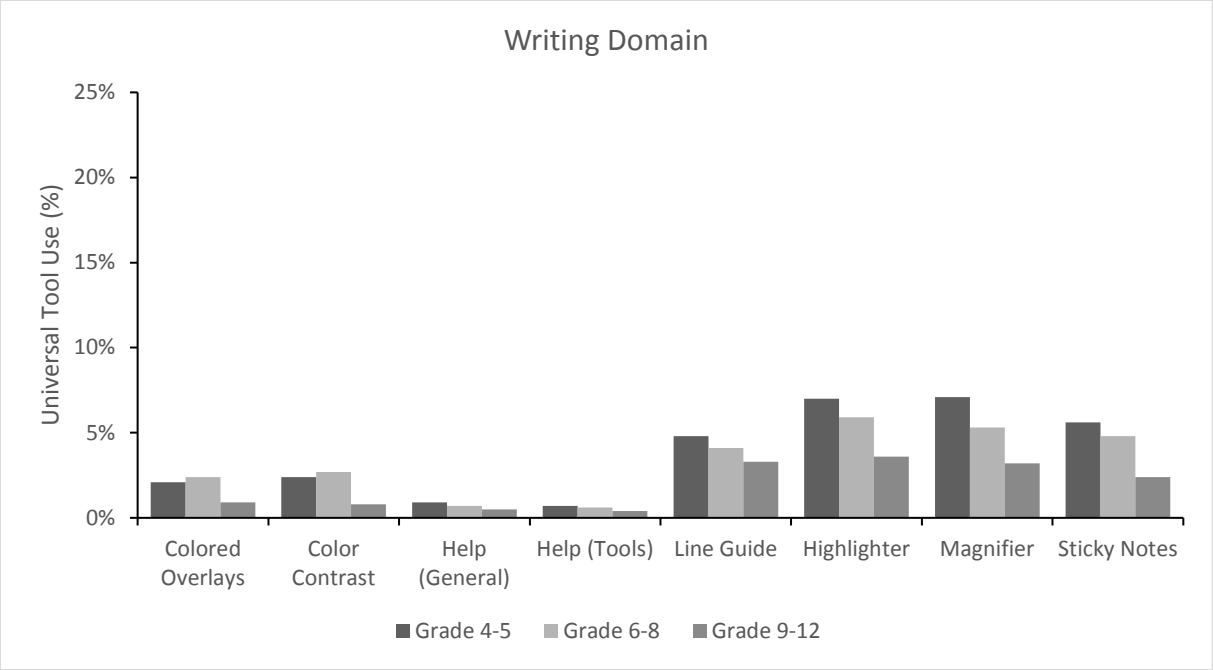

Figure 6. Frequency of universal tool use across grade-level clusters in the writing domain

In conclusion, Grades 6-8 students had the highest tool activation rate in the listening section, whereas in the reading domain, a higher percentage of Grades 4-5 students activated the universal tools. With respect to the speaking section, Grades 4-5 and Grades 6-8 students both accessed the tools more than students in the other grades. Grades 4-5 students displayed the highest use of the Magnifier, and Grades 6-8 students showed highest use of the Line Guide and the Highlighter in the speaking section. Similarly, in the writing section, activation of the tools was highest among these two grade-level clusters, with slightly more Grades 6-8 students activating the Line Guide, Highlighter, Magnifier, and Sticky Notes than students in the other grade-level clusters. Across all domains, Grade 1 students activated all the universal tools at a lower percentage rate than students in the other grade-level clusters. The grade-level cluster breakdown of the findings for ELs also showed that tool activation decreased among all grades as students progressed on the test from the listening to the speaking or writing domains, with the

exception of Grades 4-5 students in the reading domain, where relatively high percentage of students accessed each tool.

4.2. Findings from Research Question 2

Findings related to Research Question 2 (To what extent do ELs with and without disabilities differ in their use of the accessibility features embedded in an ELP assessment?) explore the differences in tool use between ELs with and without IEPs by grade-level cluster and by disability type. It also discusses how ELs with 504 Plans use tools.

4.2.1. Difference in Tool Use between ELs with and without IEPs

The second research question pertains to the difference in the use of universal tools between ELs with IEPs and ELs without IEPs. This section presents the frequency and descriptive statistics of tool activation by ELs with and without IEPs. Table 8 shows ELs' tool use across all four domains, and Table 9 summarizes the tool activation of ELs with IEPs across all four domains. Figures 7-10 report findings for each domain. In addition to frequency and descriptive statistics, the findings of group differences using Mann Whitney U tests (due to the violations of normality and homogeneity) and effect sizes (r) are reported in Tables 10-13.

Table 8. Frequency and descriptive statistics for universal tool activation across all domains among ELs

	Listening (*n*= 956,051)					Reading (*n*=955,163)					Speaking (*n*=938,822)					Writing (*n*=485,252)				
Universal tools	Use %	Ran.	$\bar{X}$	Med.	sd	Use%	Ran.	$\bar{X}$	Med.	sd	Use %	Ran.	$\bar{X}$	Med.	sd	Use %	Ran.	$\bar{X}$	Med.	sd
Color Overlay	2.4	1-17	1.68	1.0	1.098	2.0	1-15	1.65	1.0	1.100	1.4	1-19	1.51	1.0	0.944	1.7	1-19	1.53	1.0	1.000
Color Contrast	2.2	1-19	1.78	1.0	1.224	2.0	1-16	1.78	1.0	1.249	1.4	1-9	1.58	1.0	1.082	1.9	1-21	1.68	1.0	1.160
Help (General)	2.9	1-22	1.12	1.0	0.450	0.8	1-9	1.13	1.0	0.460	1.0	1-10	1.1	1.0	0.393	0.7	1-9	1.1	1.0	0.389
Help (Tools)	1.8	1-10	1.11	1.0	0.432	0.5	1-11	1.17	1.0	0.558	0.7	1-11	1.15	1.0	0.530	0.5	1-9	1.15	1.0	0.497
Line Guide	8.5	1-111	1.91	1.0	1.900	7.7	1-93	2.9	2.0	3.502	4.4	1-192	1.88	1.0	2.046	4.1	1-113	1.63	1.0	1.543
Highlighter	4.5	1-963	12.06	5.0	23.251	10.0	1-973	17.33	8.0	28.767	3.7	1-951	11.58	5.0	26.152	5.3	1-1191	12.07	6.0	24.915
Magnifier	9.0	1-240	21.84	2.0	39.041	9.1	1-295	24.77	2.0	43.665	5.1	0-236	48.08	2.0	64.933	4.9	1-464	16.82	2.0	26.713
Sticky Notes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4.1	1-117	2.01	1.0	2.255

Table 9. Frequency and descriptive statistics for universal tool activation across all domains among ELs with IEPs

	Listening (*n*=147,065)					Reading (*n*=146,882)					Speaking (*n*=143,536)					Writing (*n*=92,756)				
Universal tools	Use %	Ran.	$\bar{X}$	Med.	sd	Use%	Ran.	$\bar{X}$	Med.	sd	Use %	Ran.	$\bar{X}$	Med.	sd	Use %	Ran.	$\bar{X}$	Med.	sd
Color Overlay	3.2	1-19	1.73	1.0	1.209	2.4	1-14	1.65	1.0	1.077	1.8	1-21	1.53	1.0	1.052	1.9	1-10	1.49	1.0	0.9000
Color Contrast	2.9	1-16	1.85	1.0	1.353	2.4	1-29	1.84	1.0	1.377	1.9	1-19	1.61	1.0	1.095	2.0	1-10	1.64	1.0	1.110
Help (General)	2.8	1-9	1.13	1.0	0.476	1.0	1-6	1.13	1.0	0.435	1.0	1-6	1.11	1.0	0.393	0.8	1-5	1.09	1.0	0.370
Help (Tools)	2.0	1-10	1.12	1.0	0.451	0.7	1-7	1.2	1.0	0.595	0.9	1-5	1.15	1.0	0.467	0.6	1-13	1.18	1.0	0.731
Line Guide	10.1	1-870	2.07	1.0	7.402	8.6	1-311	2.83	2.0	4.429	5.1	1-41	1.98	1.0	2.015	4.2	1-53	1.61	1.0	1.575
Highlighter	5.9	1-968	14.22	6.0	26.524	11.4	1-475	18.37	9.0	30.289	4.7	1-1,150	13.79	6.0	31.73	5.5	1-453	12.87	6.0	23.153
Magnifier	11.8	1-202	16.25	2.0	34.15	11.0	1-188	19.99	2.0	39.120	6.1	1-199	36.15	2.0	57.401	5.4	1-236	14.51	2.0	24.831
Sticky Notes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4.4	1-60	2.08	1.0	2.588

The tool activation rate by ELs without IEPs was very similar to or slightly lower than the activation by the entire EL group (findings from Research Question 1, which combined data from ELs with and without IEPs). Due to the similarity of the data from ELs without IEPs to the data from the overall EL group, the findings below focus on tool activation by ELs with IEPs. A comparison of tool activation between ELs with and without IEP is provided when relevant.

It must be noted that the sample size of the ELs with IEPs group was much smaller than the sample size of EL without IEPs. Approximately 11% of the entire EL group was identified as having an IEP. However, percentages of students activating the universal tools are always higher among the ELs with IEPs than it was for ELs without IEPs across all domains. The differences in the percentages were more noticeable in the listening and reading domains than in the speaking and writing domains, especially for the Line Guide, Highlighter, and Magnifier (Figures 7-10).

In the listening domain (Figure 7), the Magnifier (11.8%, *Med.* = 2) and Line Guide (10.1%, *Med.* = 1) were the tools most commonly accessed by ELs with IEPs. Their activation rate of the Magnifier and Line Guide tools was 2.8% and 1.6% respectively higher than ELs'. The Highlighter was the third most activated tool among ELs with IEPs, as it attracted 5.9% (*Med.* = 6) of the students. On the other hand, 4.5% (*Med.* = 5) of ELs without IEPs used the tool (i.e., there was a 1.4% difference in activation rate between the groups). Activation of the Color tools was less than other tools for both groups, yet a slightly higher percentage of ELs with IEPs accessed Color Overlay (.8% difference) and Color Contrast (.7% difference) than ELs without IEPs. Similarly, Help was the least accessed feature among both groups, with a similar percentage of ELs with and without IEPs activating those tools.

Additionally, Mann Whitney U tests were run for the listening domain to determine if there were significant differences in the universal tool activation between ELs with and without

IEPs. The two groups significantly differed with respect to their activation of the Color Overlay, Color Contrast, Line Guide, Highlighter, and Magnifier (see Table 10). Activation of the Color Overlay tool by ELs with IEPs (mean rank = 14,117.29) was statistically significantly higher than that of ELs without IEPs (mean rank = 13,829.06), U = 55,426,225.50, z = 2.547, p = 0.011; but the effect size was small (r = 0.02). Similarly, for Color Contrast, activation of the tool was significantly higher among ELs with IEPs (mean rank = 12,991.30) than ELs without IEPs (mean rank = 12,748.13), U = 46,144,675, z = 2.173, p = 0.030; but the effect size was trivial (r = 0.01). The tests also showed the two groups were significantly different with respect to activation of the Line Guide tool: ELs with IEPs had a significantly higher activation rate (mean rank = 49,276.53) than ELs without IEPs (mean rank = 48,128.71), U = 623,963,911, z = 5.142, p = 0.000; but with small effect size (r = 0.02). Activation of the Highlighter was also statistically significantly higher for ELs with IEPs (mean rank = 27,213.38) than ELs without IEPs (mean rank = 25,436), U = 197,236,889, z = 10.161, p = 0.000; but the effect size was negligible (r = 0.04). Unlike the other tools, the Magnifier proved to be more appealing to ELs without IEPs than to ELs with IEPs. Tests showed statistically significant higher Magnifier tool use among ELs (mean rank = 52,040.35) than among ELs with IEPs (mean rank = 50,567.18). The group effect was significant (U = 725,890,114.50, z = -6.216, p = 0.000), but the effect size was small suggesting a trivial difference (r = -0.02). In sum, significant group effects were found for the Color Overlay, Color Contrast, Line Guide, Highlighter, and Magnifier tools, which could be attributed to the large sample size. However, the effect sizes for all the tools were near zero, suggesting that the differences were not meaningful.

Table 10. Mann Whitney U results in the listening domain

	Total *n*	*U*	*z*	*p*-value	*r*
Color Overlay	27,755	55,426,225.50	2.547	0.011	0.02
Color Contrast	25,576	46,144,675.00	2.173	0.030	0.01
Line Guide	96,611	623,963,911.00	5.142	0.000	0.02
Highlighter	51,466	197,236,889.00	10.161	0.000	0.04
Magnifier	103,587	725,890,114.50	-6.216	0.000	-0.02

In the reading domain (Figure 8), ELs with IEPs generally activated the tools more than ELs, except for the Highlighter and Magnifier. For ELs with IEPs, activation of all the tools in the reading domain decreased, compared to activation in the listening domain, except for the Highlighter. The Highlighter was the most accessed tool in the reading domain, with almost twice as many ELs with and without IEPs activating the tool in the reading section than in the listening section; slightly more ELs with IEPs (11.4%; *Med.* = 9) used the tool than ELs (10%; *Med.* = 8), resulting in a 1.4% difference in Highlighter activation between the two groups. There was also about a 2% difference in Magnifier activation between the two groups, with slightly more ELs with IEPs activating the tool (11%, *Med.* = 2) than ELs (9.1 %, *Med.* = 2). The Line Guide was the third most activated tool, and about 1% more ELs with IEPs (8.6%, *Med.* = 2) accessed it than ELs without IEPs (7.7%, *Med.* = 2). The two Color tools and two Help tools were the least commonly activated features for both groups. The difference in the activation of these tools between ELs with and without IEPs was also smaller than it was for other tools. While .4% more ELs with IEPs activated the two Color tools than ELs without IEPs, the use of two Help tools was nearly the same (.2% difference between the groups).

Mann Whitney U findings for the reading domain (Table 11) showed a significant group effect for four of the universal tools: Color Contrast, Line Guide, Highlighter, and Magnifier. ELs with IEPs (mean rank = 11,626.98) had a statistically higher activation of Color Contrast than ELs without IEPs (mean rank = 11,411.18), U = 34,843,407.50, z = 1.983, p = 0.047.

However, the effect size was small (r = 0.01). Similarly, for the Line Guide, ELs with IEPs (mean rank = 43,235.24) had statistically significantly high activation rates in comparison to ELs without IEPs (mean rank = 42,513.88), U = 457,603,848.50, z = -3.183, p = 0.001; but with small effect size (r = -0.01). For the Highlighter, ELs with IEPs (mean rank = 56,954.12) had a statistically higher activation of the tool than ELs without IEPs (mean rank = 55,850.02), U = 813,289,852.5, z = 4.079, p = 0.000; but with negligible effect size (r = 0.01). Finally, there was significant group difference for activation of the Magnifier tool (U = 686,896,719.5, z = -5.298, p = 0.000), with a significantly higher activation rate among ELs without IEPs (mean rank = 51,882.80) than among ELs with IEPs (mean rank = 50, 567.33). Yet, the effect size was small (r = -0.02). In conclusion, significant group differences were found in tool activation in the reading domain, but the effect size was negligible, similar to the listening domain.

Table 11. Mann Whitney U results in the reading domain

	Total *n*	*U*	*z*	*p*-value	*r*
Color Contrast	22888	34,843,407.50	1.983	0.047	0.01
Line Guide	86258	457,603,848.50	-3.183	0.001	-0.01
Highlighter	112029	813,289,852,5	4.079	0.000	0.01
Magnifier	103353	686,896,719.50	-5.298	0.000	-0.02

Activation of all tools by ELs with and without IEPs, except for the two Help tools, decreased even more in the speaking domain (Figure 9) than it did from the listening and the reading domain. ELs with IEPs had a higher tool activation rate than ELs without IEPs in general, but the group difference narrowed in the speaking domain compared to the listening and reading domains. The Magnifier was the most activated tool, with 1% more ELs with IEPs (6.1%, *Med.* = 2) than ELs without IEPs (5.1%, *Med.* = 2) accessing the tool. For the second most accessed tool, the Line Guide, .7% more ELs with IEPs (5.1%, *Med.* = 1) activated the tool than ELs without IEPs (4.4%, *Med.* = 1), indicating small group differences. Similarly, 1% more ELs with IEPs (4.7%, *Med.* = 6) activated the Highlighter than ELs without IEPs (3.7%,

Med. = 5). The two Color and Help tools did not appeal much to either group. 1.4% of ELs without IEPs and about 2% of ELs with IEPs accessed the Color tools. Activation of two Help tools was about 1% among both groups.

In the speaking domain, two groups were significantly different from each other with respect to activation of three of the universal tools: the Line Guide, Highlighter, and Magnifier, the Mann Whitney U findings show (Table 12). Regarding the Line Guide, ELs with IEPs (mean rank = 25, 017.60, *Med.* = 1) demonstrated statistically higher activation of the tools than ELs without IEPs (mean rank = 24,317.32), U = 155,997,573.50, z = 4.444, p = 0.000; but with a small effect size (r = 0.02). Similarly, for the Highlighter, ELs with IEPs (mean rank = 21.693.23) showed a statistically higher activation rate of the tool than ELs without IEPs (mean rank = 20, 595.27), U = 123,554,972.50, z = 6.907, p = 0.000, while the group effect was small (r = 0.03). In contrast, ELs without IEPs (mean ranks =28,865.06) had a significantly higher activation of the Magnifier tool than ELs with IEPs (mean rank = 26,668.85), U = 195,639,703.50, z = -11.841, p = 0.000; but the group effect was negligible (r = -0.05). In short, although we observed statistical group difference between ELs without and ELs with IEPs, the effect sizes were negligible, suggesting the group effect was not meaningful in the speaking domain either.

Table 12. Mann Whitney U results in the speaking domain

	Total *n*	*U*	*z*	*p*-value	*r*
Line Guide	48843	155,997,573.50	4.444	0.000	0.02
Highlighter	41546	123,554,972.50	6.907	0.000	0.03
Magnifier	57053	195,639,703.50	-11.841	0.000	-0.05

ELs' activation of all the tools in the writing domain (Figure 10) was also lower than in the listening and reading domains. Additionally, the difference in the percentage of ELs without and ELs with IEPs using the tools was the smallest in the writing domain, with a 0.2% difference between the groups for activation of the Highlighter, the most accessed tool of the domain (ELs

with IEPs: 5.5%, *Med*. =6; ELs without IEPs: 5.3%, *Med*. = 6). Magnifier activation was about the same: 5.4% of ELs with IEPs and 4.9% of ELs without IEPs activated the tool (0.5% difference). Sticky Notes activation was also comparatively high, relative to other tools, among ELs with IEPs (4.4%, *Med.* = 1) and ELs without IEPs (4.1%, *Med.* = 1), with a 0.3% difference between the two groups. The difference between the two groups with regard to the use of the Help and Color tools was also very small, ranging from 0.1% to 0.2%.

In the writing domain, significant differences between ELs without and ELs with IEPs occurred in Highlighter and Magnifier activation the Mann Whitney U findings show (Table 13). With respect to the Highlighter, ELs with IEPs had a significantly higher activation ate of the tool (mean rank = 15,801.99) than ELs without IEPs (mean rank = 15,353.85), U = 68,009,856.50, z = 3.303, p = 0.001; but the effect size was small (r = 0.02). On the contrary, activation of the Magnifier was statistically significantly higher among ELs without IEPs (mean rank = 14,616.29) than among ELs with IEPs (mean rank = 14,084.37), U = 58,281,918.50, z = -4.247, p = 0.000, with small effect size (r = -0.02).

Table 13. Mann Whitney U results in the writing domain

	Total *n*	*U*	*z*	*p*-value	*r*
Highlighter	30,856	68,009,856.50	3.303	0.001	0.02
Magnifier	29,047	58,281,918.50	-4.247	0.000	-0.02

In conclusion, universal tool activation was higher among ELs with IEPs than ELs without IEPs. The difference between the two groups was relatively large in the listening and reading domains compared to the speaking and writing domains. Universal tool activation by ELs without IEPs was more similar to the entire EL group than was universal tool activation by ELs with IEPs. Additionally, ELs without and ELs with IEPs differed significantly with respect to the activation of particular tools. Nevertheless, the effect sizes were small in all cases, indicating that the differences were negligible.

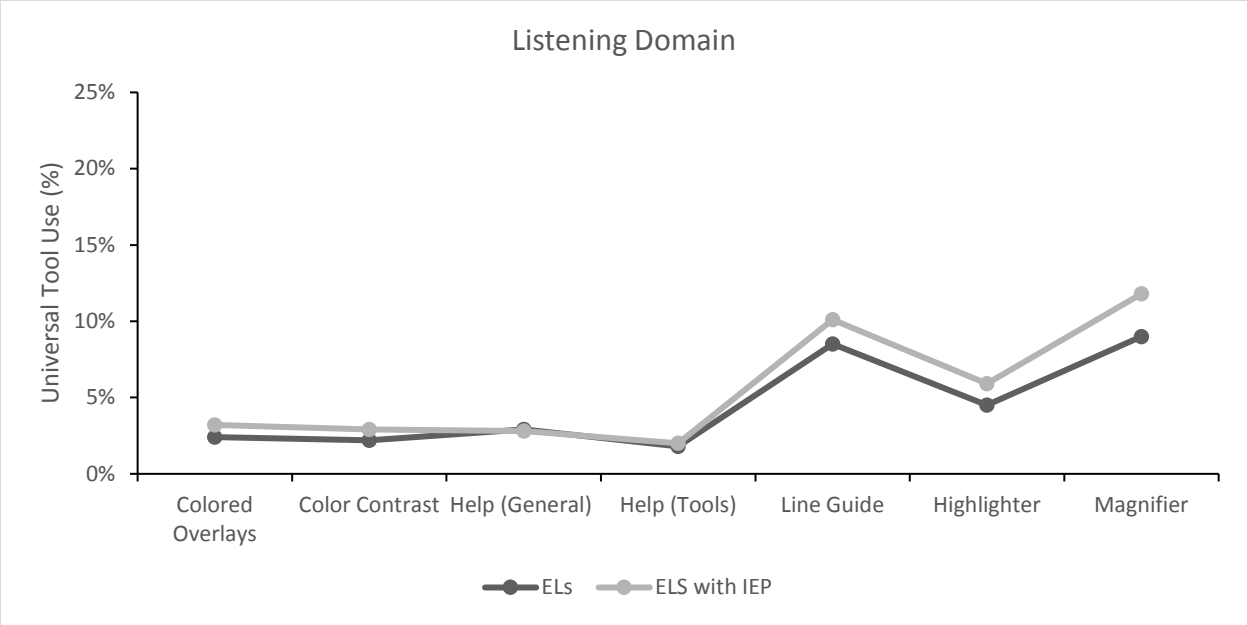

Figure 7. Frequency of universal tool use in the listening domain, ELs without vs. ELs with IEPs

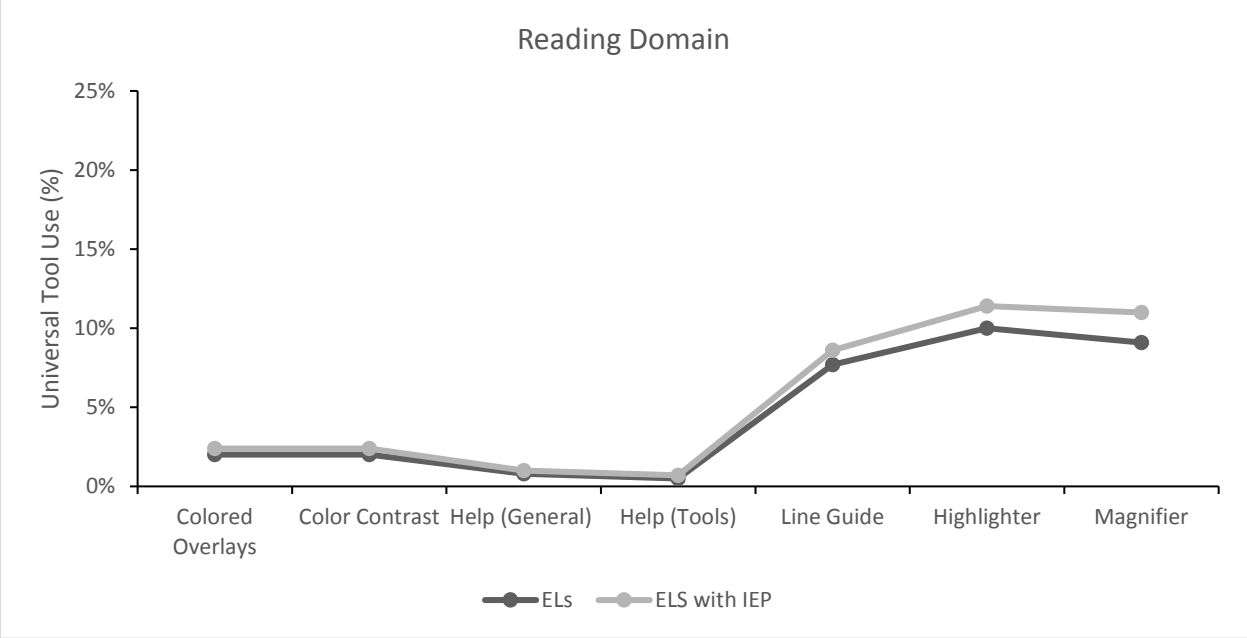

Figure 8. Frequency of universal tool use in the reading domain, ELs without vs. ELs with IEPs

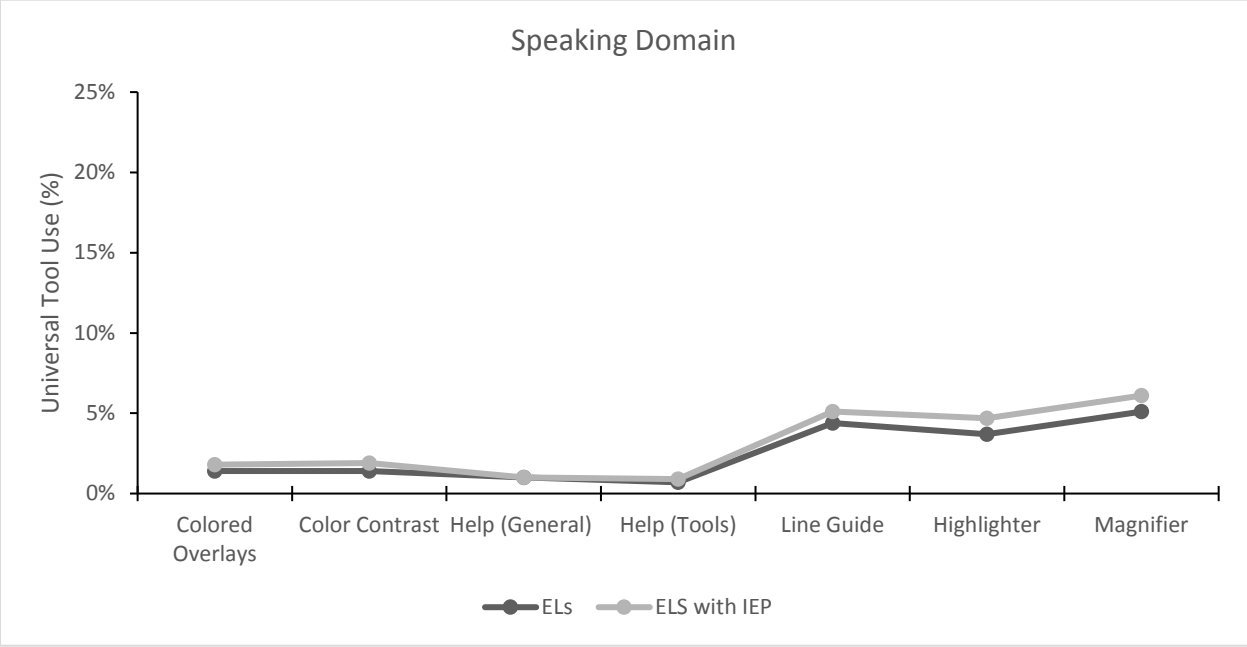

Figure 9. Frequency of tool use in the speaking domain, ELs without vs. ELs with IEPs

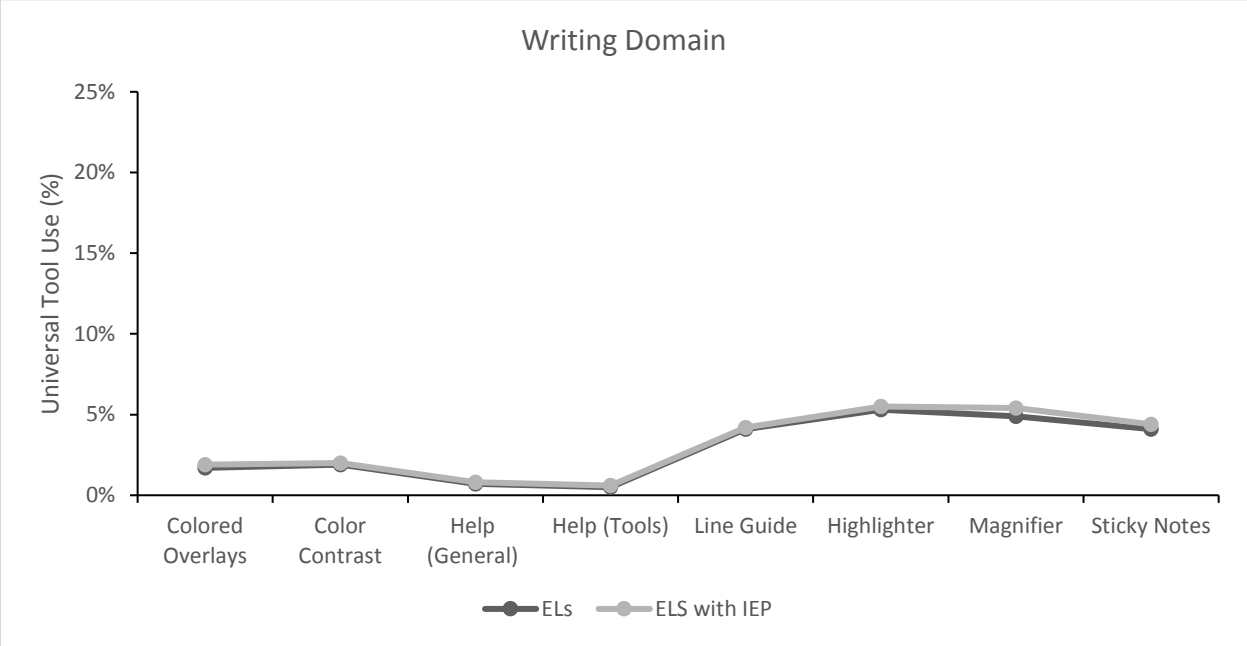

Figure 10. Frequency of tool use in the writing domain, ELs without vs. ELs with IEPs

4.2.2. Use of Tools among ELs with and without IEPs by Grade-Level Cluster

In addition to the overall comparison between ELs with and without IEPs, this study also compares ELs with and ELs without IEPs for each grade-level cluster (see Tables 14-17 and

Figures 11-15). The most salient findings with respect to grade-level cluster comparisons between the two groups are described below.

In the listening domain (Table 14), in all grades, a higher percentage of ELs with IEPs accessed the universal tools compared to ELs without IEPs. Also, Grades 6-8 students in both groups made the highest use of the tools, whereas Grade 1 students in both groups activated the tools the least. However, the differences between ELs with and ELs without IEPs were small across grade-level clusters. The largest differences occurred for the Magnifier tool. In Grades 2-3 (Figure 12), Magnifier activation by ELs with IEPs (10.0%) was 1.5% higher than by ELs without IEPs (8.5%). Grades 4-5 (Figure 13) and Grades 6-8 (Figure 14) ELs with IEPs activated the Magnifier at a slightly higher rate than ELs without IEPs in those grade clusters, resulting in about a 2% difference between the two groups. The highest difference in Magnifier use between ELs with and ELs without IEPs was among Grades 9-12 students, with 2.6% more ELs with IEPs than those without accessing the tool (Figure 15).

Tool activation among ELs with and ELs without IEPs in each grade-level cluster was similar in the reading domain (Figure 15) to what it was in the listening domain. At most, there was about 1% difference between ELs with and ELs without IEPs in each grade-level cluster. In both groups, the Line Guide, Highlighter, and Magnifier activation was highest among Grades 4-5 students (Figure 13). For both groups, activation of two Color tools was high in the Grades 4-5 and Grades 6-8 (Figure 14) clusters. The tools were the least activated by Grade 1 students (Figure 11), both ELs with and ELs without IEPs. Yet, the median activation rate of the Magnifier tool for both Grade 1 ELs without IEPs (*Med.* = 5) and Grade 1 ELs with IEPs (*Med.* = 4) was higher than the median activation rate for all other grade-level clusters (*Med.* = 1).

In the speaking domain (Table 16), tool activation was again very similar between ELs with and ELs without IEPs in each grade-level cluster, with a slightly higher percentage of ELs with IEPs making use of the tools. Specifically, the group differences across grade-level clusters were less than 1%. Grades 6-8 students, both ELs with and ELs without IEPs, accessed the tools at a higher rate than did students in the other grade–level clusters. One exception was activation of the Magnifier tool. Magnifier access was slightly higher among Grades 4-5 students in both groups. Universal tool activation was again less common among Grade 1 ELs, both those with and without IEPs. However, the median Magnifier activation rate was very large for Grade 1 in both groups (ELs without IEPs: Med. = 110, ELs with IEPs: Med. = 55).

In the writing domain (Table 17), the difference between ELs with and without IEPs across grade-level clusters was less than 0.5%. Grade 4-5 ELs with and without IEPs accessed the tools at a higher rate than did students in other grades. The percentage of Grades 9-12 students accessing the tools, on the other hand, was lowest among all the grade clusters in both groups.

In conclusion, the comparison between ELs with and without IEPs in each grade-level cluster did not yield considerable differences. In the listening domain, comparatively large differences between the two groups were observed, with the most difference in activation of the Magnifier. In the reading domain, the difference in tool access between ELs with and ELs without IEPs across the grade-level clusters was about 1%. In the speaking and writing domains, the difference between the two groups in any given grade-level cluster was less than 1%.

Table 14. Frequency and descriptive statistics for universal tool use among ELs with and ELs without IEPs by grade-level cluster in the listening domain

		Color Overlay					Color Contrast					Help (General)					Help (Tools)				
		Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
ELs	**Grade 1** (*n* = 138,422)	0.6	1-16	1.79	1.0	1.323	0.7	1-15	1.94	1.0	1.531	2.0	1-7	1.13	1.0	.429	0.8	1-5	1.1	1.0	.393
	Grades 2-3 (*n* = 282,508)	1.5	1-12	1.76	1.0	1.169	1.5	1-19	1.82	1.0	1.249	2.6	1-8	1.1	1.0	.369	1.3	1-9	1.1	1.0	.409
	Grades 4-5 (*n* = 160,216)	2.6	1-17	1.74	1.0	1.163	2.5	1-17	1.82	1.0	1.247	2.9	1-7	1.1	1.0	.379	1.8	1-8	1.11	1.0	.432
	Grades 6-8 (*n* = 174,290)	5.0	1-12	1.70	1.0	1.093	4.6	1-14	1.84	1.0	1.26	3.9	1-11	1.1	1.0	.457	3.0	1-9	1.13	1.0	.482
	Grades 9-12 (*n* = 200,615)	2.6	1-16	1.52	1.0	.924	2.0	1-13	1.55	1.0	.965	3.0	1-22	1.12	1.0	.575	2.2	1-10	1.08	1.0	.396
ELs with IEPs	**Grades 1** (*n* = 12,787)	0.8	1-8	1.77	1.0	1.219	0.9	1-13	2.12	1.0	1.839	1.8	1-8	1.22	1.0	.741	.8	1-7	1.21	1.0	.736
	Grades 2-3 (*n* = 31,230)	1.6	1-11	1.79	1.0	1.354	1.5	1-16	2.0	1.0	1.615	2.6	1-7	1.17	1.0	.545	1.4	1-4	1.11	1.0	.414
	Grades 4-5 (*n* = 31,495)	2.6	1-13	1.77	1.0	1.23	2.7	1-14	1.89	1.0	1.398	2.6	1-9	1.12	1.0	.465	1.7	1-10	1.15	1.0	.593
	Grades 6-8 (*n* = 40,360)	5.3	1-19	1.79	1.0	1.27	4.8	1-16	1.88	1.0	1.334	3.5	1-6	1.13	1.0	.455	3.0	1-6	1.12	1.0	.434
	Grades 9-12 (*n* = 31,193)	3.7	1-8	1.58	1.0	.977	2.8	1-9	1.63	1.0	1.072	2.6	1-4	1.07	1.0	.315	2.3	1-3	1.07	1.0	.279

		Line Guide					Highlighter					Magnifier				
		Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
ELs	**Grade 1** (*n* = 138,422)	4.3	1-111	1.71	1.0	2.307	2.2	1-579	12.38	6.0	24.053	5.5	1-240	42.65	3.0	47.541
	Grades 2-3 (*n* = 282,508)	6.2	1-66	1.91	1.0	1.867	4.2	1-377	10.07	5.0	16.796	8.5	1-219	26.91	2.0	42.225
	Grades 4-5 (*n* = 160,216)	8.8	1-42	1.94	1.0	1.791	5.5	1-252	9.74	5.0	15.841	11.9	1-188	19.35	2.0	37.179
	Grades 6-8 (*n* = 174,290)	12.4	1-98	1.99	1.0	2.031	6.9	1-963	14.93	6.0	30.385	12.9	1-239	13.59	2.0	31.651
	Grades 9-12 (*n* = 200,615)	11.2	1-65	1.86	1.0	1.73	3.6	1-637	13.3	5.0	25.347	6.6	1-168	18.25	2.0	35.922
ELs with IEPs	**Grade 1** (*n* = 12,787)	5.2	1-21	1.89	1.0	2.222	2.8	1-243	15.55	7.0	27.689	6.4	1-202	38.92	3.0	46.395
	Grades 2-3 (*n* = 31,230)	7.0	1-20	2.01	1.0	1.877	4.9	1-389	12.41	6.0	20.529	10.0	1-160	24.18	2.0	40.267
	Grades 4-5 (*n* = 31,495)	9.5	1-26	2.01	1.0	1.787	6.8	1-462	11.87	6.0	21.52	13.9	1-168	15.80	2.0	33.827
	Grades 6-8 (*n* = 40,360)	13.5	1-86	2.12	1.0	2.448	8.0	1-968	16.54	7.0	32.971	15.2	1-162	11.79	2.0	29.314
	Grades 9-12 (*n* = 31,193)	11.6	1-870	2.1	1.0	14.537	4.4	1-256	14.10	7.0	21.675	9.2	1-132	11.47	2.0	28.405

Table 15. Frequency and descriptive statistics for universal tool use among ELs with and ELs without IEPs by grade-level cluster in the reading domain

		Color Overlay					Color Contrast					Help (General)					Help (Tools)				
		Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
ELs	**Grade 1** (*n* = 138,372)	0.6	1-9	1.58	1.0	1.005	0.6	1-9	1.67	1.0	1.078	0.9	1-8	1.15	1.0	.524	0.4	1-3	1.10	1.0	.334
	Grades 2-3 (*n* = 282,384	1.5	1-15	1.68	1.0	1.145	1.6	1-14	1.83	1.0	1.308	0.8	1-9	1.14	1.0	.485	0.4	1-8	1.16	1.0	.557
	Grades 4-5 (*n* = 160,149)	3.2	1-13	1.75	1.0	1.197	3.4	1-16	1.92	1.0	1.375	1.1	1-5	1.13	1.0	.436	0.8	1-7	1.22	1.0	.614
	Grades 6-8 (*n* = 174,202)	3.4	1-14	1.62	1.0	1.050	3.5	1-12	1.74	1.0	1.179	0.9	1-7	1.12	1.0	.436	0.7	1-11	1.20	1.0	.606
	Grades 9-12 (*n* = 200,056)	1.4	1-15	1.50	1.0	0.950	1.3	1-12	1.58	1.0	1.023	0.4	1-6	1.1	1.0	.375	0.3	1-6	1.11	1.0	.471
ELs with IEPs	**Grade 1** (*n* = 12,784)	0.7	1-11	1.99	1.0	1.739	0.7	1-10	1.90	1.5	1.392	1.0	1-3	1.17	1.0	.448	0.4	1-2	1.09	1.0	.293
	Grades 2-3 (*n* = 31,202)	1.6	1-7	1.64	1.0	1.038	1.5	1-9	1.83	2.0	1.096	1.2	1-5	1.14	1.0	.454	0.6	1-5	1.22	1.0	.558
	Grades 4-5 (*n* = 31,459)	3.2	1-11	1.75	1.0	1.195	3.3	1-15	1.93	1.0	1.483	1.4	1-5	1.15	1.0	.461	1.0	1-5	1.22	1.0	.626
	Grades 6-8 (*n* = 40,344	3.4	1-14	1.61	1.0	1.022	3.5	1-29	1.85	1.0	1.477	0.9	1-6	1.1	1.0	.404	0.8	1-7	1.21	1.0	.68
	Grades 9-12 (*n* = 31,093)	1.9	1-7	1.51	1.0	0.850	1.6	1-8	1.59	1.0	1.007	0.4	1-3	1.11	1.0	.342	0.4	1-3	1.15	1.0	.381

		Line Guide					Highlighter					Magnifier				
		Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
ELs	**Grade 1** (*n* = 138,372)	4.3	1-66	2.26	1.0	3.226	3.3	1-568	13.82	6.0	23.681	5.4	1-164	49.01	5.0	53.817
	Grades 2-3 (*n* = 282,384	7.6	1-59	3.37	2.0	4.053	8.5	1-692	14.76	7.0	23.798	9.0	1-198	28.82	2.0	46.490
	Grades 4-5 (*n* = 160,149)	13.5	1-90	3.36	2.0	3.771	16.8	1-704	16.74	9.0	26.148	16.1	1-22	17.6	2.0	37.243
	Grades 6-8 (*n* = 174,202)	8.6	1-57	2.33	1.0	2.55	13.9	1-973	18.83	9.0	31.411	10.2	1-295	18.47	2.0	38.507
	Grades 9-12 (*n* = 200,056)	4.8	1-93	2.08	1.0	2.492	7.8	1-674	20.98	9.0	35.729	5.3	1-160	26.01	2.0	43.905
ELs with IEPs	**Grade 1** (*n* = 12,784)	5.1	1-40	2.36	1.0	3.945	3.7	1-207	16.58	8.0	24.039	6.1	1-168	47.01	4.0	52.876
	Grades 2-3 (*n* = 31,202)	8.0	1-39	3.15	2.0	3.924	8.2	1-348	16.98	8.0	27.034	10.0	1-178	27.62	2.0	44.950
	Grades 4-5 (*n* = 31,459)	13.4	1-43	3.32	2.0	3.958	17.3	1-470	17.73	9.0	28.865	17.7	1-188	15.44	2.0	34.495
	Grades 6-8 (*n* = 40,344	9.2	1-311	2.47	2.0	5.732	14.1	1-475	19.13	9.0	32.837	11.4	1-158	16.34	2.0	35.324
	Grades 9-12 (*n* = 31,093)	5.1	1-37	2.09	1.0	2.414	8.2	1-338	19.74	9.0	31.328	6.3	1-30	19.01	2.0	37.917

Table 16. Frequency and descriptive statistics for universal tool use among ELs with and ELs without IEPs by grade-level cluster in the speaking domain

		Color Overlay					Color Contrast					Help (General)					Help (Tools)				
		Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med*	*sd*
ELs	**Grade 1** (*n* = 136,101)	0.5	1-14	1.59	1.0	1.031	0.5	1-13	1.71	1.0	1.254	0.7	1-10	1.10	1.0	.451	0.4	1-6	1.10	1.0	.468
	Grades 2-3 (*n* = 279,478)	0.9	1-19	1.54	1.0	1.014	1.0	1-19	1.61	1.0	1.142	0.6	1-6	1.11	1.0	.408	0.4	1-9	1.14	1.0	.521
	Grades 4-5 (*n* = 158,322)	1.9	1-13	1.53	1.0	.957	2.1	1-16	1.62	1.0	1.142	1.0	1-6	1.11	1.0	.396	0.8	1-7	1.17	1.0	.558
	Grades 6-8 (*n* = 171,303)	2.7	1-13	1.52	1.0	.928	2.7	1-16	1.56	1.0	.964	0.8	1-6	1.10	1.0	.374	1.5	1-10	1.17	1.0	.58
	Grades 9-12 (*n* = 193,618)	1.2	1-8	1.45	1.0	.845	1.0	1-17	1.48	1.0	1.085	0.9	1-6	1.09	1.0	.376	0.8	1-6	1.12	1.0	.434
ELs with IEPs	**Grade 1** (*n* = 12,435)	0.5	1-5	1.55	1.0	.829	0.6	1-5	1.6	1.0	.867	0.6	1-3	1.15	1.0	.428	0.6	1-3	1.16	1.0	.463
	Grades 2-3 (*n* = 30,817)	1.1	1-11	1.58	1.0	1.178	1.1	1-8	1.73	1.0	1.168	0.7	1-4	1.12	1.0	.404	0.5	1-3	1.11	1.0	.368
	Grades 4-5 (*n* = 31,092)	2.0	1-21	1.6	1.0	1.267	2.3	1-19	1.54	1.0	1.264	1.0	1-6	1.12	1.0	.434	.9	1-5	1.20	1.0	.577
	Grades 6-8 (*n* = 49,592)	2.8	1-11	1.52	1.0	.983	3.0	1-15	1.58	1.0	1.025	0.6	1-4	1.11	1.0	.363	1.4	1-4	1.14	1.0	.436
	Grades 9-12 (*n* = 29,600)	1.6	1-6	1.43	1.0	.791	1.4	1-6	1.51	1.0	.915	0.8	1-4	1.11	1.0	.394	0.7	1-5	1.13	1.0	.447

		Line Guide					Highlighter					Magnifier				
		Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Max.	$\bar{X}$	*Med*	*sd*
ELs	**Grade 1** (*n* = 136,101)	2.9	1-90	1.64	1.0	2.134	2.0	1-251	10.14	5.0	15.153	4.0	1-210	80.76	110	72.898
	Grades 2-3 (*n* = 279,478)	3.7	1-28	1.89	1.0	1.781	3.4	1-399	8.71	5.0	14.873	5.0	1-218	55.40	2.0	67.799
	Grades 4-5 (*n* = 158,322)	5.7	1-192	1.99	1.0	2.748	5.0	1-669	9.01	4.0	18.312	7.1	1-184	34.95	2.0	55.08
	Grades 6-8 (*n* = 171,303)	6.2	1-44	1.95	1.0	1.845	5.5	1-685	14.07	6.0	29.923	6.4	1-236	37.95	2.0	63.05
	Grades 9-12 (*n* = 193,618)	4.0	1-38	1.74	1.0	1.554	2.6	1-951	17.12	6.0	43.279	3.4	1-208	44.76	2.0	58.621
ELs with IEPs	**Grade 1** (*n* = 12,435)	3.0	1-41	1.8	1.0	2.521	2.0	1-212	15.34	7.0	23.989	4.3	1-190	66.47	52	66.388
	Grades 2-3 (*n* = 30,817)	4.0	1-23	1.91	1.0	1.935	3.8	1-432	11.78	6.0	22.366	5.7	1-199	47.13	2.0	62.495
	Grades 4-5 (*n* = 31,092)	6.0	1-32	2.12	1.0	2.096	5.9	1-360	10.35	5.0	20.265	7.9	1-166	29.48	2.0	50.959
	Grades 6-8 (*n* = 49,592)	6.7	1-31	2.0	1.0	1.997	6.2	1-1,150	15.47	6.0	37.942	7.1	1-173	30.86	2.0	56.916
	Grades 9-12 (*n* = 29,600)	3.9	1-39	1.83	1.0	1.797	3.5	1-500	17.77	6.0	40.83	4.1	1-166	32.64	2.0	52.194

Table 17. Frequency and descriptive statistics for universal tool use among ELs with and ELs without IEPs by grade-level cluster in the writing domain

		Color Overlay					Color Contrast					Help (General)					Help (Tools)				
		Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med*	*sd*	Use %	Ran.	$\bar{X}$	*Med*	*sd*
ELs	**Grades 4-5** (*n* = 116,155)	2.1	1-17	1.67	1.0	1.199	2.4	1-18	1.84	1.0	1.354	0.9	1-5	1.1	1.0	.362	0.7	1-5	1.18	1.0	.54
	Grades 6-8 (*n* = 172,213)	2.5	1-19	1.51	1.0	0.950	2.8	1-21	1.65	1.0	1.105	0.8	1-5	1.09	1.0	.331	0.6	1-9	1.17	1.0	.544
	Grades 9-12 (*n* = 196,884)	.9	1-11	1.39	1.0	0.766	0.8	1-11	1.48	1.0	0.887	0.5	1-9	1.1	1.0	.471	0.4	1-4	1.08	1.0	.351
ELS with IEPs	**Grades 4-5** (*n* = 22,337)	1.9	1-9	1.56	1.0	0.984	2.2	1-9	1.74	1.0	1.235	1.2	1-4	1.09	1.0	.34	0.8	1-5	1.19	1.0	.558
	Grades 6-8 (*n* = 39,876)	2.4	1-9	1.5	1.0	0.905	2.7	1-9	1.65	1.0	1.087	0.8	1-5	1.1	1.0	.399	0.7	1-13	1.2	1.0	.91
	Grades 9-12 (*n* = 30,543)	1.3	1-10	1.4	1.0	0.778	1.1	1-10	1.48	1.0	0.974	0.5	1-4	1.09	1.0	.358	0.4	1-4	1.13	1.0	.488

		Line Guide					Highlighter					Magnifier					Sticky Notes				
		Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
ELs	**Grades 4-5** (*n* = 116,155)	4.8	1-15	1.66	1.0	1.268	7.0	1-463	10.41	6.0	16.072	7.2	1-350	16.82	2.0	29.114	5.6	1-25	2.04	1.0	1.912
	Grades 6-8 (*n* = 172,213)	4.3	1-113	1.66	1.0	1.915	6.0	1-677	13.5	7.0	26.293	5.4	1-219	15.79	2.0	24.61	5.0	1-117	2.05	1.0	2.666
	Grades 9-12 (*n* = 196,884)	3.4	1-41	1.59	1.0	1.277	3.7	1-1191	11.88	5.0	30.330	3.3	1-464	18.28	2.0	26.313	2.2	1-43	1.9	1.0	1.853
ELS with IEPs	**Grades 4-5** (*n* = 22,337)	5.1	1-53	1.64	1.0	1.948	7.3	1-453	11.79	6.0	21.319	7.7	1-172	13.31	2.0	25.336	5.7	1-60	2.18	1.0	2.965
	Grades 6-8 (*n* = 39,876)	4.2	1-47	1.61	1.0	1.565	5.7	1-305	13.31	7.0	2.281	5.5	1-149	15.54	2.0	24.91	4.9	1-41	2.13	1.0	2.586
	Grades 9-12 (*n* = 30,543)	3.5	1-12	1.57	1.0	1.048	4.0	1-446	13.51	6.0	28.29	3.8	1-236	14.32	2.0	23.838	2.8	1-21	1.82	1.0	1.887

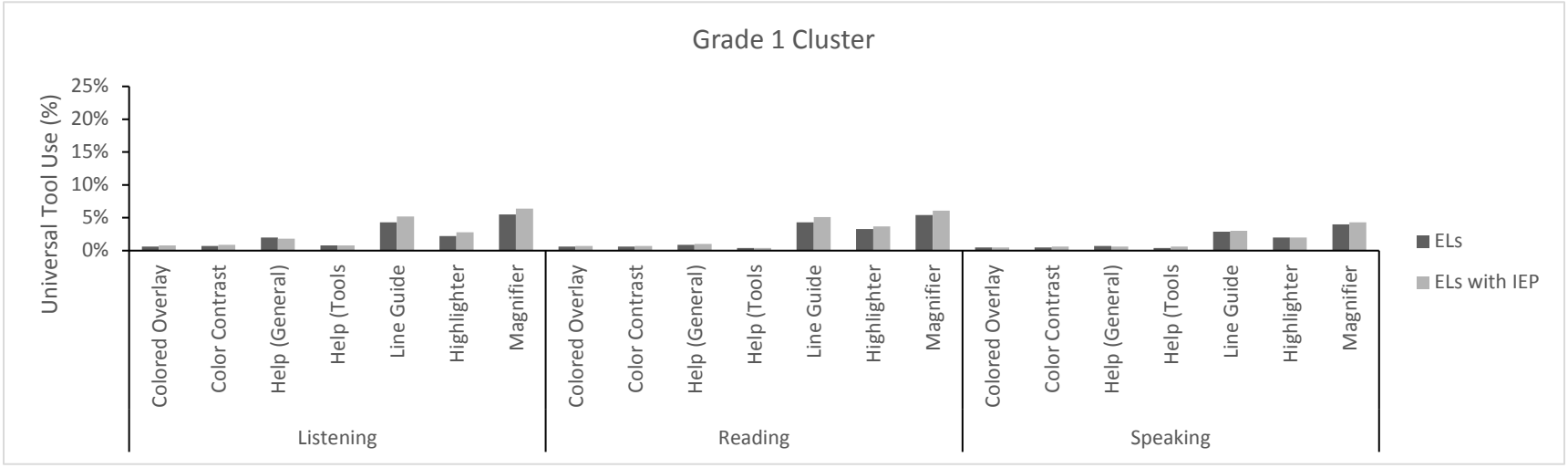

Figure 11. Frequency of universal tool use in Grade 1, ELs with vs. ELs without IEPs

Figure 12. Frequency of universal tool activation in Grades 2-3, ELs with vs. ELs without IEPs

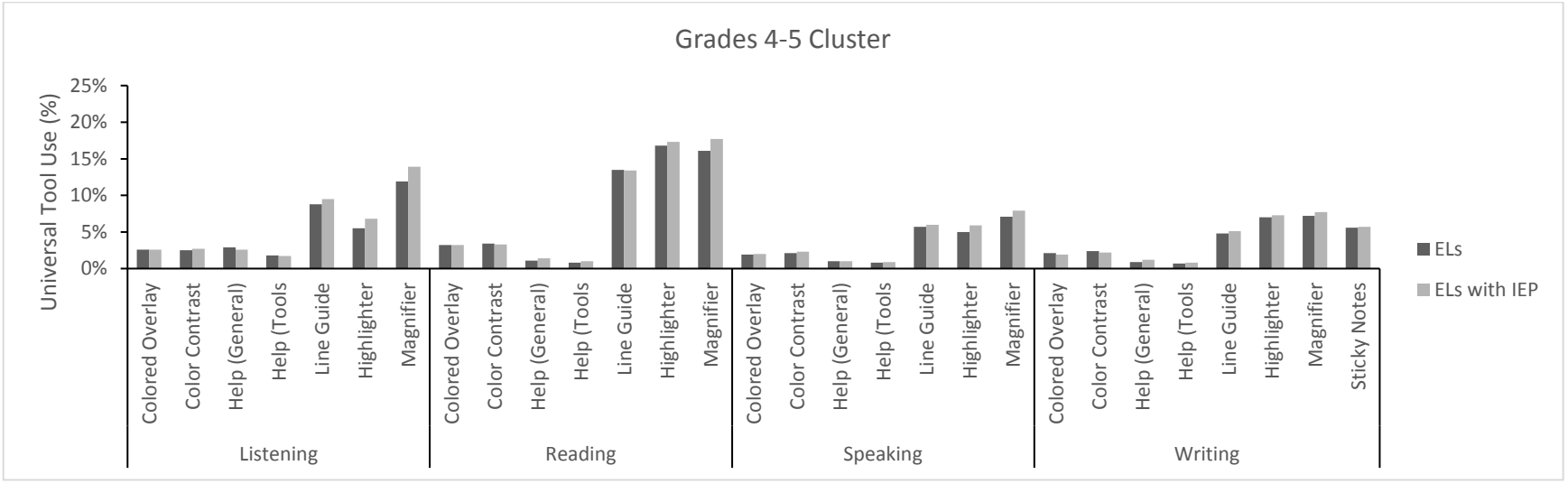

Figure 13. Frequency of universal tool use in Grades 4-5, ELs with vs. ELs without IEPs

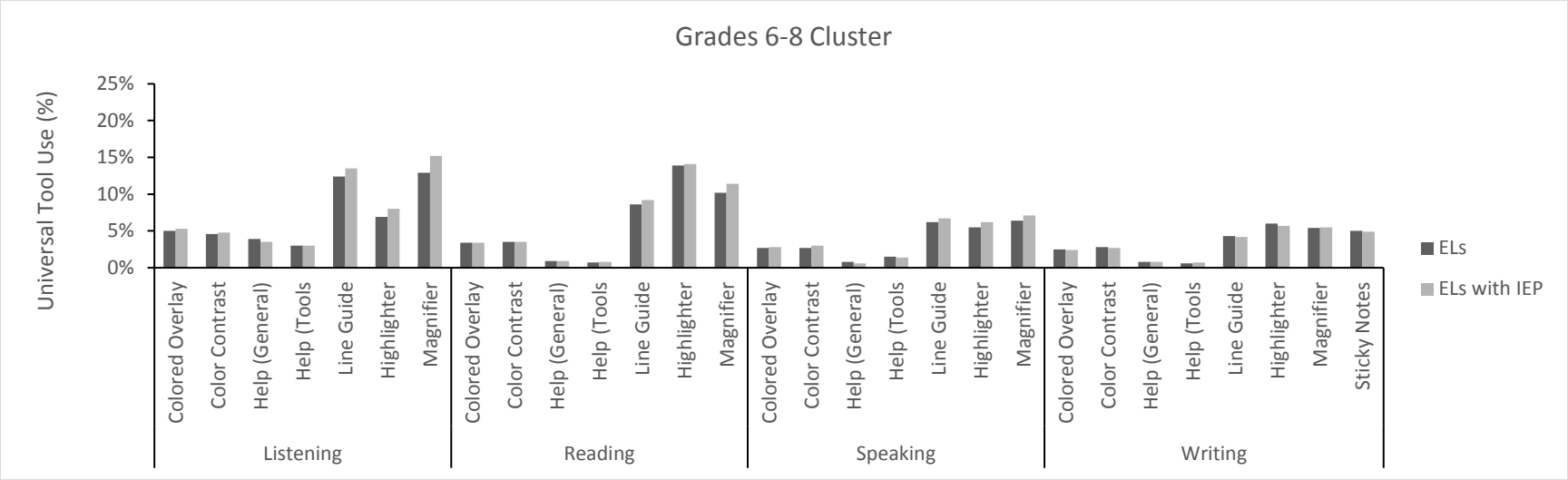

Figure 14. Frequency of universal tool use in Grades 6-8, ELs with vs. ELs without IEPs

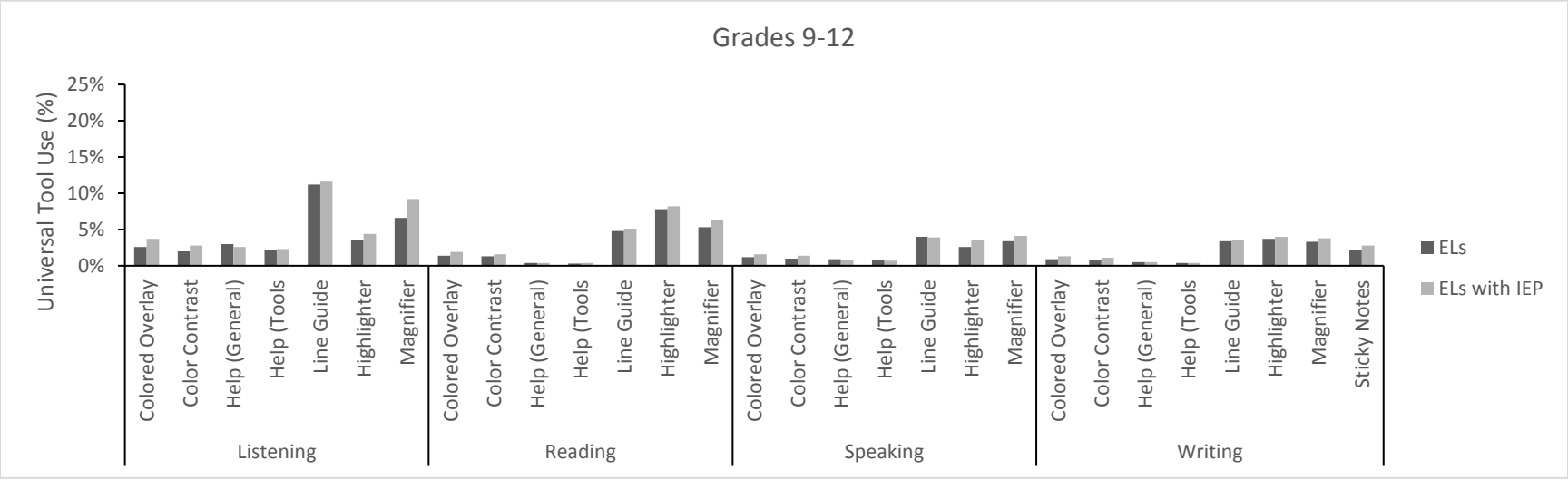

Figure 15. Frequency of universal tool use in Grades 9-12, ELs with vs. ELs without IEPs

4.2.3. Use of Tools among ELs with IEPs by Disability Type

Universal tool activation was postulated to vary across different disability types for ELs with IEPs. The frequency of universal tool activation by disability type is reported in Tables 18-21. Although information about disability type is missing (marked as NA in Tables 16-19) for the majority of the students, tool activation across disability types is detailed based on the available data: autism spectrum disorder, deaf-blindness, developmental delay, hearing impairment, intellectual disability, multiple disability, other health impairment, orthopedic impairment, serious emotional disability, specific learning disability, speech or language impairment, traumatic brain injury, and visual impairment.

In all domains, specific learning disability (18%; e.g., disorder in the basic psychological processes involved in understanding or using languages, such as dyslexia and developmental aphasia) was the most common primary disability type, and speech or language impairment (6.5%) was the second most common disability type. Deaf-blindness (0.01%) and visual impairments (0.05%) were the least common primary disability types. (For the writing domain, percentages slightly differ. See Appendix A for the distribution of disability types in each domain.) However, disability type was not reported for about 66% of the students.

In the listening domain (Table 18), the activation rate of all tools was highest among ELs with serious emotional disability. Students with other health impairment had the second highest rate of tool activation. One exception was the Magnifier: 42% ELs with visual impairment activated the Magnifier, and 19% of ELs with deaf-blindness activated the tool. Meanwhile, students with deaf-blindness, developmental delay, and visual impairments had generally low tool activation rates. Specifically, ELs with deaf-blindness did not activate any tools other than

the Line Guide and Magnifier. Similarly, activation of the Color Overlay, Help, and Highlighter tools was low among ELs with visual impairments.

In the reading domain (Table 19), ELs with visual impairment and serious emotional disability generally displayed high tool activation rates. In particular, Magnifier access was highest among ELs with visual impairments (51.2%). In addition, ELs with other health impairments displayed high activation of the Line Guide, Highlighter, and Magnifier tools. ELs with deaf-blindness displayed the highest activation rate of the Highlighter tool. However, these students did not activate the Color and Help tools, and their activation rate of the Line Guide and Magnifier was very low. Also, students with developmental delay and intellectual disability generally had low rates of tool use.

In the speaking domain (Table 20), tool activation was generally higher among ELs with serious emotional disability than among other disability groups. The ELs with visual impairment group was an exception, in that a high percentage of students in this group activated the Color Contrast and Magnifier. Yet, in comparison to the reading domain, Magnifier activation dropped sharply in the speaking domain. Likewise, ELs with deaf-blindness showed comparatively high activation of the Color Overlay and Line Guide. However, they did not activate any other tools in the speaking section. The universal tool activation in the speaking section was lowest among ELs with developmental delay, intellectual disability, and traumatic brain injury.

Finally, in the writing domain (Table 21), as in the other domains, ELs identified with serious emotional disability and visual impairments had comparatively high rate of tool activation. In particular, the rate of Magnifier access was high among ELs with visual impairments, as half of those students activated this tool in the writing section. There were also some exceptions to this general trend. Students with deaf-blindness displayed a high rate of

activation of the Line Guide, Highlighter, and Sticky Notes, but they did not access the other tools. Students with speech or language impairment also displayed comparatively high rates of Highlighter activation. Meanwhile, universal tool activation in the writing domain was generally the lowest among ELs with intellectual disability.

In conclusion, ELs with serious emotional disability made comparatively high use of the tools. Similarly, ELs with Visual Impairments had a higher rate of activation of the Color Contrast and Magnifier tools. However, students with developmental delays, intellectual disabilities, and traumatic brain injury displayed low universal tool activation. In the listening domain, deaf-blind students displayed a high rate of Magnifier activation. Similarly, in the reading domain, students with visual impairments accessed the Magnifier more frequently than did other students. In the speaking and writing domains, a high percentage of deaf-blind students activated the Line Guide.

Table 18. Universal tool use by disability type in the listening domain

Disability Type	n	Color Overlay Use %	Color Contrast Use %	Help (General) Use %	Help (Tools) Use %	Line Guide Use %	Highlighter Use %	Magnifier Use %
Autism Spectrum Disorder	1,913	2.4	2.0	4.0	2.6	9.8	6.4	11.5
Deaf-blindness	16	0	0	0	0	6.3	0	18.8
Developmental Delay	2,979	1.2	1.2	1.6	1.0	6.1	3.7	7.7
Hearing impairment, including deafness	461	3.0	3.0	2.6	1.7	9.8	3.9	11.7
Intellectual Disability	1,860	2.4	1.7	3.2	1.8	9.8	4.9	11.5
Multiple Disability	441	2.7	.7	2.5	1.4	10.0	7.0	12.7
Other Health Impairment	4,691	4.6	4.4	3.3	2.5	13.4	8.8	15.5
Orthopedic Impairment	176	2.3	1.7	1.7	3.4	11.4	5.1	9.7
Serious Emotional Disability	1,281	5.7	5.2	4.3	3.7	17.3	9.8	16.7
Specific Learning Disability	26,599	3.3	2.8	2.6	1.9	9.9	5.7	11.8
Speech or Language Impairment	9,576	2.0	2.0	2.4	1.5	7.8	4.9	9.6
Traumatic Brain Injury	123	0	.8	3.3	0	6.5	7.3	13.8
Visual Impairment, including blindness	82	1.2	3.7	0	0	14.6	1.2	41.5
NA	97,583	3.3	3.0	2.8	2.1	10.3	5.9	11.9

Table 19. Universal tool use by disability type in the reading domain

Disability Type	n	Color Overlay Use %	Color Contrast Use %	Help (General) Use %	Help (Tools) Use %	Line Guide Use %	Highlighter Use %	Magnifier Use %
Autism Spectrum Disorder	1,908	1.9	2.1	1.5	.9	8.6	10.5	9.6
Deaf-blindness	18	0	0	0	0	5.6	22.2	5.6
Developmental Delay	2,970	1.3	1.2	1.0	.5	6.9	5.6	7.7
Hearing impairment, including deafness	483	2.3	1.4	.4	1.0	10.6	12.6	9.9
Intellectual Disability	1,844	1.7	1.4	.9	.5	6.7	9.2	10.0
Multiple Disability	438	1.6	1.1	.5	.9	8.0	9.4	11.9
Other Health Impairment	4,673	3.0	3.3	1.3	1.0	11.1	13.2	13.3
Orthopedic Impairment	176	2.3	2.3	1.1	.6	13.1	12.5	11.9
Serious Emotional Disability	1,260	3.3	3.3	1.3	1.0	10.6	13.1	12.0
Specific Learning Disability	26,520	2.5	2.6	.9	.6	9.2	12.2	11.7
Speech or Language Impairment	9,560	1.8	1.9	1.0	.6	8.6	10.0	11.1
Traumatic Brain Injury	121	.8	.8	.8	.8	8.3	11.6	14.0
Visual Impairment, including blindness	80	3.8	3.8	1.3	0	11.3	8.8	51.2
NA	97,544	2.5	2.4	1.0	.7	8.4	11.5	10.9

Table 20. Universal tool use by disability type in the speaking domain

Disability Type	n	Color Overlay Use %	Color Contrast Use %	Help (General) Use %	Help (Tools) Use %	Line Guide Use %	Highlighter Use %	Magnifier Use %
Autism Spectrum Disorder	1,838	1.5	1.7	1.4	1.4	5.8	5.8	6.3
Deaf-blindness	11	9.1	0.0	0.0	0.0	9.1	0.0	0.0
Developmental Delay	2,915	0.7	0.8	0.4	0.4	3.3	2.5	4.7
Hearing impairment, including deafness	446	2.0	1.6	1.1	1.3	4.0	5.4	5.8
Intellectual Disability	1,758	1.3	1.0	0.7	1.0	3.8	3.5	4.7
Multiple Disability	425	0.9	0.9	0.7	0.7	5.9	5.2	8.2
Other Health Impairment	4,568	2.5	3.0	1.5	1.2	6.4	6.0	7.5
Orthopedic Impairment	173	0.6	1.7	1.7	1.7	4.6	2.9	3.5
Serious Emotional Disability	1,196	3.9	3.7	1.9	1.7	8.7	9.9	9.7
Specific Learning Disability	25,917	1.9	1.9	1.1	0.9	5.0	4.7	5.9
Speech or Language Impairment	9,409	1.4	1.5	1.0	0.8	5.0	4.1	5.7
Traumatic Brain Injury	118	0.8	0.8	0.0	0.0	4.2	1.7	3.4
Visual Impairment, including blindness	75	1.3	4.0	2.7	1.3	4.0	4.0	18.7
NA	94,687	1.8	1.9	1.0	0.8	5.1	4.7	6.2

Table 21. Universal tool use by disability type in the writing domain

Disability Type	n	Color Overlay Use %	Color Contrast Use %	Help (General) Use %	Help (Tools) Use %	Line Guide Use %	Highlighter Use %	Magnifier Use %	Sticky Notes Use %
Autism Spectrum Disorder	922	1.2	1.7	1.7	1.0	3.9	6.8	6.0	4.8
Deaf-blindness	9	0.0	0.0	0.0	0.0	22.2	11.1	11.1	11.1
Developmental Delay	97	0.0	2.1	1.0	0.0	8.2	10.3	14.4	4.1
Hearing impairment, including deafness	302	2.3	2.0	1.0	0.7	5.0	4.3	6.6	3.3
Intellectual Disability	1,374	0.6	0.5	0.4	0.4	3.6	4.0	4.4	2.5
Multiple Disability	296	1.0	2.0	1.0	0.3	4.7	5.1	6.4	2.4
Other Health Impairment	3,103	2.8	3.1	0.9	0.7	4.9	6.1	5.6	5.5
Orthopedic Impairment	95	1.1	1.1	0.0	1.1	6.3	6.3	3.2	4.2
Serious Emotional Disability	930	2.3	3.0	1.0	1.1	5.8	5.5	6.8	5.8
Specific Learning Disability	19,911	1.7	1.7	0.8	0.6	3.6	5.0	5.1	4.0
Speech or Language Impairment	3,295	2.0	2.1	1.3	0.8	5.7	7.1	6.7	5.8
Traumatic Brain Injury	81	0.0	1.2	1.2	0.0	3.7	3.7	7.4	4.9
Visual Impairment, including blindness	49	2.0	6.1	2.0	0.0	6.1	2.0	51.0	0.0
NA	62,292	2.0	2.1	0.8	0.6	4.2	5.6	5.4	4.4

4.2.4. Use of Tools among ELs with 504 Plans

In addition to ELs with IEPs, we conducted a separate analysis of ELs with 504 Plans (Table 22). The majority of ELs with 504 Plans were not identified as having IEPs. Their pattern of accessing universal tools was similar to that of ELs without any identified disability. ELs with 504 Plans was a relatively small sample of students: approximately 8,000 students in the listening, reading, and speaking domains and about 4,750 in the writing domain.

Overall, universal tool activation among ELs with 504 Plans was similar to that of ELs without IEPs. The differences in the rate of tool access between two groups were small across all domains. The largest difference occurred for the Line Guide in the listening domain. About 1% more ELs with 504 Plans accessed the Line Guide than ELs. Similarly, in the writing domain, Highlighter and Sticky Notes activation by ELs with 504 Plans was about 1% higher than by ELs.

In comparison to ELs with IEPs, tool activation by ELs with 504 Plans was generally lower, except for use of the Highlighter and Sticky Notes tools in the writing domain. About 1% more ELs with 504 Plans accessed these tools. Additionally, the differences between the two groups were particularly small in the speaking and writing domains. The largest difference between ELs with 504 Plans and ELs with IEPs was in activation of the Magnifier tool in the listening and reading domains, about a 2% difference between the two groups. Finally, in the writing domain, a higher percentage of ELs with 504 Plans accessed the Highlighter (6.1%), and that rate was higher than that of both ELs (5.3%) and ELs with IEPs (5.5%).

Table 22. Frequency and descriptive statistics for universal tool use among ELs with 504 Plans across all domains

	Listening (n = 8,340)				Reading (n = 8,327)				Speaking (n = 8,121)				Writing (n = 4,753)			
Universal tools	Use %	$\bar{X}$	Med.	sd	Use%	$\bar{X}$	Med.	sd	Use %	$\bar{X}$	Med.	sd	Use %	$\bar{X}$	Med.	sd
Color Overlay	2.4	1.69	1.0	1.143	2.0	1.63	1.0	0.949	1.4	1.69	1.0	1.18	1.6	1.95	1.0	1.782
Color Contrast	2.4	1.69	1.0	1.095	2.1	1.75	1.0	1.162	1.6	1.69	1.0	1.553	2.2	1.74	1.0	1.298
Help (General)	2.8	1.11	1.0	0.378	0.9	1.09	1.0	0.286	1.0	1.14	1.0	.381	0.7	1.15	1.0	0.442
Help (Tools)	1.9	1.11	1.0	0.366	0.6	1.31	1.0	0.969	0.7	1.10	1.0	0.360	0.5	1.30	1.0	0.823
Line Guide	9.4	1.91	1.0	1.667	7.6	2.50	1.0	2.937	4.9	2.03	1.0	1.922	4.1	1.63	1.0	1.801
Highlighter	5.0	13.03	5.0	22.018	9.9	16.61	8.0	32.928	4.0	14.80	6.0	32.148	6.1	12.41	6.0	21.261
Magnifier	9.5	10.82	2.0	27.692	8.9	13.21	2.0	30.903	4.8	23.77	2.0	49.791	5.0	8.84	2.0	19.228
Sticky Notes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	5.3	2.25	1.0	2.601

4.3. Findings from Research Question 3

This section discusses use of tools across proficiency levels and by proficiency level and grade-level cluster to answer Research Question 3 (To what extent do ELs at different proficiency levels vary in their use of the accessibility features embedded in an ELP assessment?)

4.3.1. Use of Tools across Different Proficiency Levels

Universal tool use was inspected across proficiency levels (PLs) in anticipation of a possible relationship between tool activation and ELP (e.g., a higher percentage of advanced learners making use of some tools like Sticky Notes). ELs were split into three groups based on their PL for each of the four domains: (a) Beginner (ELs with PLs 1-2); (b) Intermediate (ELs with PLs 3-4); and (c) Advanced (ELs with PLs 5-6). The sample size of the PL groups fluctuated across domains. Specifically, Advanced students outnumbered Beginner and Intermediate students in the listening domain. In the reading domain, there were more Beginner students than Intermediate and Advanced students. The number of Intermediate learners was highest in both the writing and speaking domains. Also, in these domains, the number of Advanced students was relatively low in comparison to that population in the listening and reading domains. Below, we elaborate on the tool activation patterns across the three PL groups

in each domain. The comparison among the three levels starts with the most accessed tools of each domain.

In the listening domain (Figure 16 and Table 23), Intermediate and Advanced students generally displayed a higher rate of tool access than did Beginners. The Line Guide and Help (General) tools were exceptions, with a slightly higher percentage of Beginners (10%) activating these tools than both Intermediate (7.8%) and Advanced (7.8%) students. Overall, differences in the percentages of overall tool access across PLs were small. In particular, tool activation was only slightly higher among students in the Intermediate group than among those in the Advanced group. For instance, 9.6% of Intermediate ELs, 9.4% of Advanced ELs, and 8.2% of Beginner ELs activated the Magnifier. Similarly, the Highlighter was accessed by 5% of Intermediate ELs, 4.6% of Advanced ELs, and 4.3% of Beginner ELs. In line with this finding, the medians for all tool activation were the same across all PLs except those for Highlighter (Beginner and Intermediate: *Med.* = 6, Advanced: *Med.* = 5).

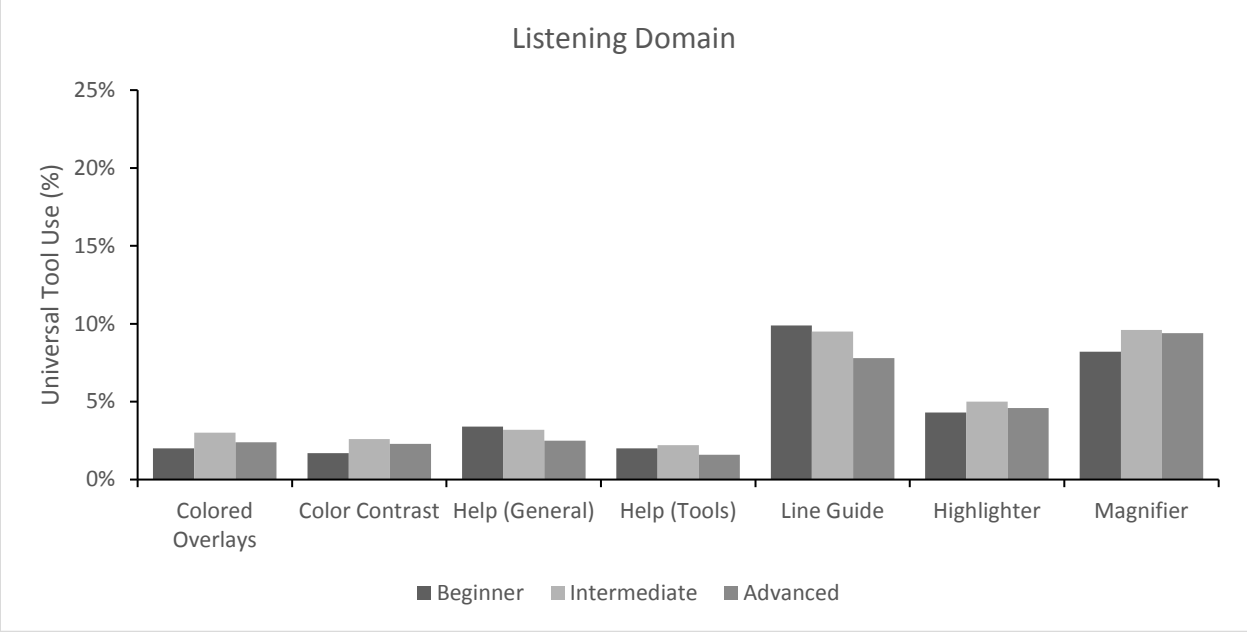

Figure 16. Frequency of universal tool use across PLs in the listening domain

Table 23. Frequency and descriptive statistics for universal tool use across different PLs in the listening domain

	Line Guide					Highlighter					Magnifier				
	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
Beginner (*n* = 227,450)	9.9	1-870	2.08	1.0	6.340	4.3	1-968	13.75	6.0	25.677	8.2	1-192	17.79	2.0	33.702
Intermediate (*n* = 349,043)	9.5	1-32	1.89	1.0	1.637	5.0	1-690	13.26	6.0	24.694	9.6	1-170	17.54	2.0	35.741
Advanced (*n* = 685,054)	7.8	1-86	1.90	1.0	1.748	4.6	1-963	11.47	5.0	21.982	9.4	1-240	23.08	2.0	40.209

	Color Overlay					Color Contrast					Help (General)					Help (Tools)				
	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
Beginner (*n* = 227,450)	2.0	1-12	1.69	1.0	1.165	1.7	1-15	1.82	1.0	1.348	3.4	1-22	1.18	1.0	.660	2.0	1-10	1.15	1.0	.546
Intermediate (*n* = 349,043)	3.0	1-16	1.69	1.0	1.132	2.6	1-19	1.80	1.0	1.248	3.2	1-9	1.11	1.0	.397	2.2	1-10	1.10	1.0	.428
Advanced (*n* = 685,054)	2.4	1-19	1.68	1.0	1.086	2.3	1-18	1.78	1.0	1.210	2.5	1-7	1.09	1.0	.360	1.6	1-9	1.09	1.0	.387

Kruskal Wallis results (Table 24) indicated statistically significant differences across the three PLs in the activation of four tools: Help (General) (H = 54.676, *p-value* = 0.000), Line Guide (H = 11.053, *p-value* = 0.000), Magnifier (H = 114.626, *p-value* = 0.000), and Highlighter (H = 149.061, *p-value* = 0.000). With respect to Help (General), post-hoc pairwise tests showed the activation of Help (General) was significantly different between Beginner and Intermediate ELs (U = 339.910, z = 5.751 p-value = 0.000, r = 0.042), and between Beginner and Advanced ELs (U = 404. 331, z = 7.301 p-value = 0.000, r = 0.046). These findings indicate that Beginners accessed the Help (General) tool significantly more frequently than did students in the other PL groups. Beginner ELs also accessed the Line Guide significantly more than did Advanced ELs (U = 748.894, z = 3.323 p-value = 0.003, r = 0.012). In addition, pairwise comparisons for the Highlighter revealed a higher rate of activation among Beginner than among Advanced ELs (U = 1,856.572, z = 9.521 p-value = 0.000, r = 0.047); the same analysis revealed higher rate of activation among Intermediate ELs than among Advanced ELs (U = -1,726.685, z = 10.083, p-value = 0.000, r = 0.022). Likewise, activation of Magnifier was statistically different between Beginner and Advanced ELs (U = -1,609.704, z = -6.485, p-value = 0.000, r = -0.046) and Intermediate vs. Advanced ELs (U = -2, 156.624, z = -9.979, p-value = 0.000, r = -0.032), with Advanced ELs accessing the tool significantly more than others. However, all of these differences were negligible as the effect sizes were small. The significant differences were possibly due to relatively large sample sizes.

Table 24. Pairwise comparison results in the listening domain

		U	*z*	*p*-value	*r*
Beginner vs. Intermediate	Help (General)	339.910	5.751	0.000	0.042
Beginner vs. Advanced	Help (General)	404. 331	7.301	0.000	0.046
	Line Guide	748.894	3.323	0.003	0.012
	Highlighter	1,856.572	9.521	0.000	0.047
	Magnifier	-1,609.704	-6.485	0.000	-0.046
Intermediate vs. Advanced	Highlighter	-1,726.685	10.083	0.000	0.022
	Magnifier	-2, 156.624	-9.979	0.000	-0.032

In the reading domain (Table 25 and Figure 17), Intermediate and Advanced ELs displayed a higher rate of tool activation across all tools than did Beginners. The overall percentages of Intermediate and Advanced students accessing the tools were very similar. They had a particularly high activation rate of the Line Guide, Highlighter, and Magnifier. A slightly higher percentage of Advanced students accessed the Line Guide (9%) and Highlighter (11.9%) than Intermediate students (8.5% and 11.5%, respectively). Also, Magnifier access was slightly higher among Intermediate ELs (10.2%) than Advanced ELs (9.7). Beginners' Line Guide access was 6.7%, Highlighter access was 8.5%, and Magnifier access was 8.6%. Additionally, the comparison between reading and listening domains revealed that the use of tools across the three PLs decreased in the reading domain except for the Highlighter and Magnifier tools. Activation of the Highlighter doubled in all PLs. Another exception was Advanced students' use of the Line Guide. Advanced ELs were the only group increasing activation of the Line Guide in the reading domain as compared to the listening domain.

Table 25. Frequency and descriptive statistics for universal tool use across different proficiency levels in the reading domain

	Line Guide					Highlighter					Magnifier				
	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
Beginner (*n* = 615,372)	6.7	1-66	2.50	1.0	2.904	8.5	1-973	16.83	8.0	27.759	8.6	1-198	22.07	2.0	40.810
Intermediate (*n* = 330,128)	8.5	1-62	3.02	2.0	3.592	11.5	1-761	17.17	9.0	27.967	10.2	1-295	23.51	2.0	43.098
Advanced (*n* = 314,984)	9.0	1-311	3.35	2.0	4.517	11.9	1-704	18.44	9.0	31.078	9.7	1-188	27.10	2.0	45.590

	Color Overlay					Color Contrast					Help (General)					Help (Tools)				
	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
Beginner (*n* = 615,372)	1.9	1-16	1.63	1.0	1.1	1.9	1-19	1.77	1.0	1.242	1.0	1-13	1.14	1.0	.498	0.6	1-11	1.17	1.0	.576
Intermediate (*n* = 330,128)	2.2	1-19	1.66	1.0	1.134	2.3	1-29	1.81	1.0	1.316	0.8	1-9	1.12	1.0	.453	0.5	1-7	1.18	1.0	.540
Advanced (*n* = 314,984)	2.1	1-19	1.63	1.0	1.064	2.1	1-24	1.78	1.0	1.256	0.6	1-6	1.12	1.0	.424	0.4	1-6	1.20	1.0	.557

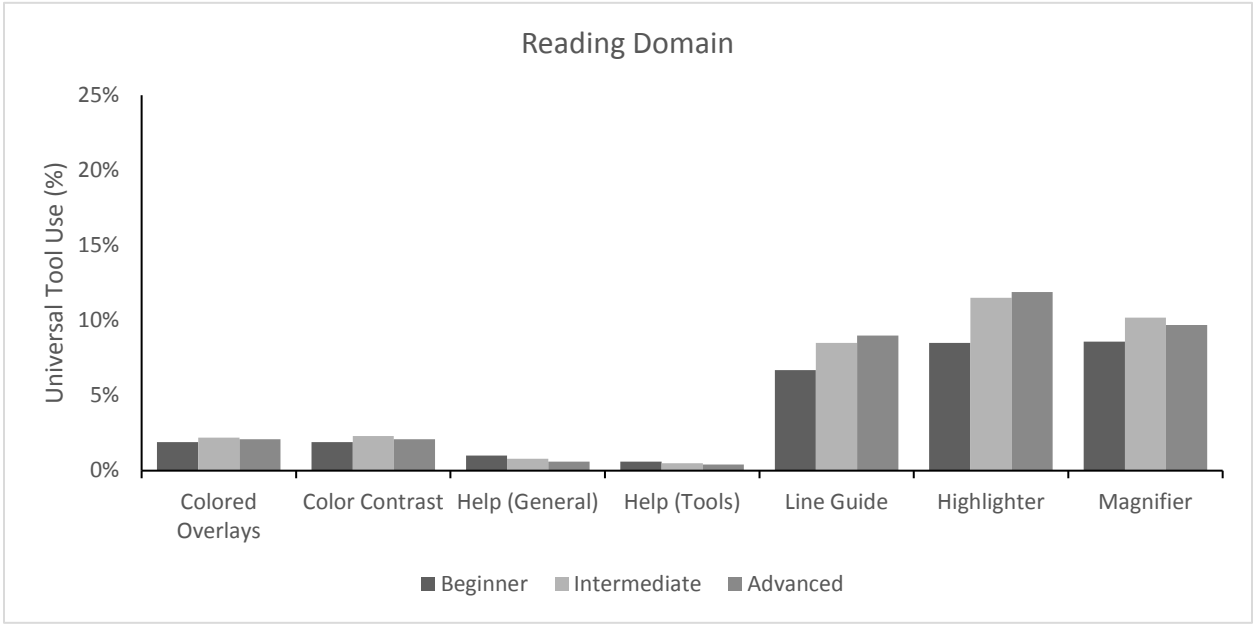

Figure 17. Frequency of universal tool use across PLs in the reading domain

With respect to significance testing and effect sizes in the reading domain, there was a statistical difference across the PLs in the activation of the Help (Tools) (H = 7.503, *p-value* = 0.023), Line Guide (H = 1,120.91, *p-value* = 0.000), Highlighter (H = 68,744, *p-value* = 0.000), and Magnifier (H = 227.198, *p-value* = 0.000).

The pairwise comparisons (Table 26) for the Help (Tools) showed that Advanced ELs accessed the tool significantly more than did Beginners (U = -95.400, z = -2.706, p-value = 0.02, r = -0.038). With respect to the Line Guide, Advanced ELs again activated the tool significantly more than Beginners (U = -6,525.796, z = -31.751, p-value = 0.000, r = -0.12), or Intermediates (U = -1,865.702, z = -8.319, p-value = 0.000, r = -0.035). In addition, Intermediates activated the tool significantly more than Beginners (U = -4,660.094, z = -22.608, p-value = 0.000, r = -0.086). However, effect sizes indicated that the only meaningful difference was between Beginner and Advanced ELs. For the Highlighter tool, all pairwise comparisons of the three PLs were significant (i.e., Beginner vs. Intermediate: U = -1,086.459, z = -4.373, p-value = 0.000, r = -0.015; Beginner vs. Advanced: U = -2,049.334, z = -8.23, p-value = 0.000, r = -0.027; Intermediate vs. Advanced: U = -962.875, z = -3.59, p-value = 0.001, r = -0.013). Specifically, Advanced ELs accessed the tool significantly more than did Beginner and Intermediate ELs, and Intermediate ELs accessed the tool significantly more than did Beginners. The Magnifier tool also yielded significant pairwise differences between all PLs (i.e., Beginner vs. Intermediate: U = -2,246.303, z = -9.83, p-value = 0.000, r = -0.033; Beginner vs. Advanced: U = -3,368.170, z = -14.292, p-value = 0.000, r = -0.05; Intermediate vs. Advanced: U = -1,121.867, z = -4.322, p-value = 0.000, r = -0.017). Again, Advanced ELs activated the tool significantly more than Intermediate and Beginner ELs, and Intermediate ELs activated the tool significantly more than

Beginners. However, the effect sizes across groups for both Highlighter and Magnifier tools were negligible and did not suggest any meaningful variation.

Table 26. Pairwise comparison results in the reading domain

		U	*z*	*p*-value	*r*
Beginner vs. Intermediate	Line Guide	-4,660.094	-22.608	0.000	-0.086
	Highlighter	-1,086.459	-4.373	0.000	-0.015
	Magnifier	-2,246.303	-9.83	0.000	-0.033
Beginner vs. Advanced	Help (Tools)	-95.400	-2.706	0.02	-0.038
	Line Guide	-6,525.796	-31.751	0.000	-0.12,
	Highlighter	-2,049.334	-8.23	0.000	-0.027
	Magnifier	-3,368.170	-14.292	0.000	-0.05
Intermediate vs. Advanced	Line Guide	-1,865.702	-8.319	0.000	-0.035
	Highlighter	-962.875	-3.59	0.001	-0.013
	Magnifier	-1,121.867	-4.322	0.000	-0.017

In the speaking domain (Table 27 and Figure 18), tool activation decreased among all PLs in comparison to the rate in the listening and reading domains. Advanced students displayed the highest activation rate of tools, whereas Beginner ELs had the lowest tool activation rate in the speaking domain. The two groups especially differed in their activation of the Line Guide, Highlighter, and Magnifier tools. For example, Line Guide activation was 6.1% among Advanced ELs and 4.3% among Beginners. There was also a 2% difference in Highlighter and Magnifier activation, with a higher percentage of Advanced ELs (5.6% and 6.8%, respectively) accessing the tool than Beginners (3.6% and 4.9%, respectively). Additionally, the overall percentage of Intermediate ELs accessing the tools was slightly higher than that of Beginners. The two Help tools were exceptions, as a higher percentage of Beginner than Intermediate and Advanced ELs activated them. However, activation of the two Help tools was very low in the speaking domain. As in the reading domain, despite the variation in percentage of use, the medians across PLs were the same for all tools except the Highlighter.

Table 27. Frequency and descriptive statistics for universal tool use across different PLs in the speaking domain

	Line Guide					Highlighter					Magnifier				
	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
Beginner (*n* = 592,363)	4.3	1-90	1.87	1.0	1.874	3.6	1-1150	13.03	5.0	29.064	4.9	1-210	41.95	2.0	59.843
Intermediate (*n* = 642,749)	4.6	1-192	1.91	1.0	2.174	4.0	1-951	10.73	5.0	24.172	5.4	1-236	48.80	2.0	66.668
Advanced (*n* = 2,690)	6.1	1-17	1.95	1.0	2.170	5.6	1-115	7.67	4.0	13.221	6.8	1-174	37.21	2.0	61.640

	Color Overlay					Color Contrast					Help (General)					Help (Tools)				
	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
Beginner (*n* = 592,363)	1.4	1-21	1.53	1.0	1.006	1.4	1-19	1.60	1.0	1.119	1.1	1-10	1.11	1.0	.440	.8	1-11	1.16	1.0	.568
Intermediate (*n* = 642,749)	1.5	1-19	1.49	1.0	.916	1.6	1-36	1.56	1.0	1.074	.9	1-6	1.09	1.0	.344	.7	1-10	1.14	1.0	.473
Advanced (*n* = 2,690)	1.6	1-6	1.44	1.0	1.033	2.1	1-5	1.65	1.0	1.009	.4	1-2	1.08	1.0	.289	.6	1-3	1.25	1.0	.577

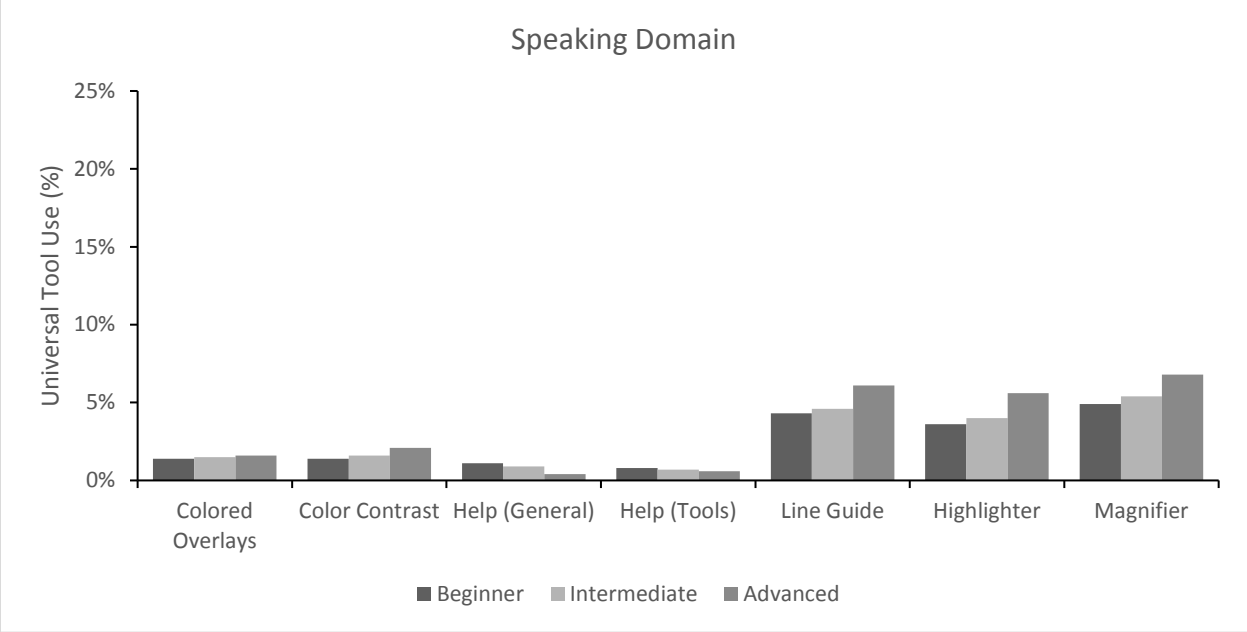

Figure 18. Frequency of universal tool use across PLs in the speaking domain

In the speaking domain, tool activation was statistically significantly different among PLs for the Help (General) (H = 6.631 *p-value* = 0.036), Line Guide (H = 6.858, *p-value* = 0.032), Highlighter (H = 135.544, *p-value* = 0.000), and Magnifier (H = 83.489, *p-value* = 0.000) tools. Pairwise comparison (Table 28) showed that Beginner and Intermediate ELs significantly differed in their activation of the Help (General), Line Guide, and Magnifier tools. While Beginners accessed Help (General) significantly more than did Intermediate ELs (U = 76.636, z = 2.575, p-value = 0.03, r = 0.024), Intermediate ELs accessed the Line Guide (U = -288.167, z = -2.4 p-value = 0.049, r = -0.01) and Magnifier (U = -1,282.142, z = -8.938, p-value = 0.000, r = -0.035) tools significantly more than did Beginners. However, effect sizes were too small to suggest meaningful differences between these two groups. For the Highlighter tool, Beginners accessed Highlighter significantly more than did Intermediate ELs (U = 1,409.421, z = 11.191, p-value = 0.000, r = 0.049) or Advanced ELs (U = 4,346.950, z = 3.891, p-value = 0.000, r = 0.026), and Intermediate ELs accessed it significantly more than did Advanced ELs (U = 2,937.529, z = 2.631, p-value = 0.009, r = 0.016). However, the effects sizes were, again, negligible.

Table 28. Pairwise comparison results in the speaking domain

		U	*z*	*p*-value	*r*
Beginner vs. Intermediate	Help (General)	76.636	2.575	0.03	0.024
	Line Guide	-288.167	-2.4	0.049	-0.01
	Highlighter	1,409.421	11.191	0.000	0.049
	Magnifier	-1,282.142	-8.938	0.000	-0.035
Beginner vs. Advanced	Highlighter	4,346.950	3.891	0.000	0.026
Intermediate vs. Advanced	Highlighter	2,937.529	2.631	0.009	0.016

Tool activation in the writing domain (Table 29 and Figure 19) was similar to that in the speaking domain. A higher percentage of Advanced students accessed all the tools, while Beginner ELs displayed the lowest activation rate, except for the Help (General) tool. There was, in particular, more variation across PLs in terms of activation of the Highlighter and Sticky Notes tools. Activation of the Highlighter was 7.3% among Advanced ELs, 5.8% among Intermediate ELs, and 3.6% among Beginners. Similarly, there was more than a 2% difference across PLs in terms of activation of the Sticky Notes tools (Advanced: 7.5%; Intermediate: 4.7%; Beginner: 2%). Additionally, unlike in other domains, the most commonly accessed features varied across PLs in the writing domain. Among the Advanced ELs, the Sticky Notes and Highlighter tools were the most accessed features. Advanced ELs' use of Sticky Notes was relatively high, and at least half of Advanced students activated Sticky Notes up to two times (*Med.* = 2). Among the Intermediate ELs, the Highlighter and Magnifier were more commonly accessed than the other features. Finally, Beginners activated the Highlighter, Magnifier, and Line Guide more than the other tools.

Table 29. Frequency and descriptive statistics for universal tool use across different PLs in the writing domain

	Color Overlay					Color Contrast					Help (General)					Help (Tools)				
	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
Beginner (*n* = 167,017)	1.3	1-19	1.64	1.0	1.158	1.4	1-21	1.61	1.0	1.155	1.1	1-9	1.11	1.0	.439	.6	1-9	1.13	1.0	.484
Intermediate (*n* = 486,852)	1.9	1-17	1.5	1.0	.923	2.1	1-15	1.68	1.0	1.134	.6	1-5	1.09	1.0	.350	.5	1-13	1.16	1.0	.558
Advanced (*n* = 4,073)	2.1	1-5	1.38	1.0	.755	2.0	1-7	1.58	1.0	1.082	.7	1-3	1.11	1.0	.416	.7	1-2	1.14	1.0	.351
	Line Guide					Highlighter					Magnifier					Sticky Notes				
	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*	Use %	Ran.	$\bar{X}$	*Med.*	*sd*
Beginner (*n* = 167,017)	3.7	1-113	1.64	1.0	1.946	3.6	1-677	12.77	5.0	28.101	3.9	1-464	17.42	2.0	28.013	2.0	1-61	1.94	1.0	2.482
Intermediate (*n* = 486,852)	4.1	1-53	1.63	1.0	.364	5.8	1-1191	12.09	6.0	24.271	5.2	1-298	16.06	2.0	25.815	4.7	1-117	2.03	1.0	2.290
Advanced (*n* = 4,073)	5.7	1-9	1.64	1.0	1.176	7.3	1-186	12.44	6.0	19.2	5.8	1-94	15.39	2.0	24.325	7.5	1-15	2.30	2.0	2.057

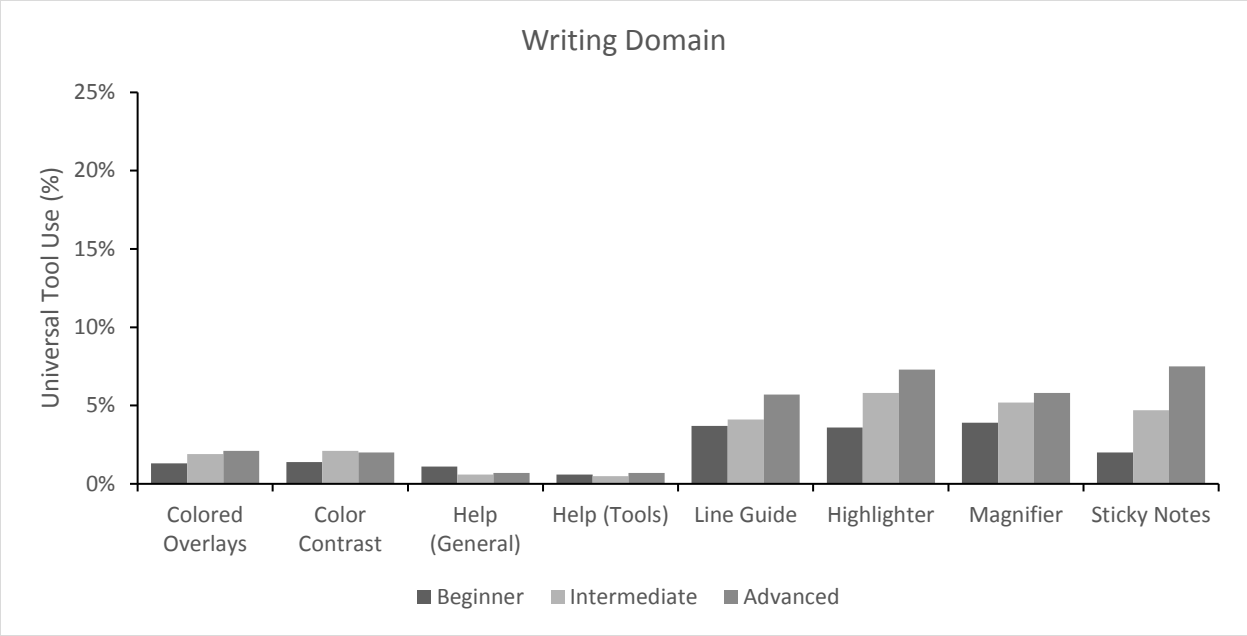

Figure 19. Frequency of universal tool use across PLs in the writing domain

In the writing domain, statistically significant differences were observed for four of the universal tools: Color Overlay (H = 7.089, *p-value* = 0.029), Color Contrast (H = 12.962, *p-value* = 0.002), Highlighter (H = 19.393, *p-value* = 0.000), and Sticky Notes (H = 48.871, *p-value* = 0.000). Pairwise comparisons (Table 28) showed Beginner and Intermediate ELs' activation of the Color Overlay (U = -169.709, z = -2.593, p-value = 0.029, r = -0.024), Color Contrast (U = -253.212, z = -3.523, p-value = 0.001, r = -0.032), and Highlighter (U = -604.551, z = -4.273, p-value = 0.000, r = -0.023) tools was statistically different in the writing domain, with Intermediate ELs accessing the tools significantly more than Beginners. However, extremely small effect sizes suggest that the difference was not meaningful. For the Sticky Notes tool, there was a statistical difference between Beginner and Intermediate (U = -730.055, z = -5.785, p-value = 0.000, r = -0.036), Beginner and Advanced (U = -2,189.457, z = -5.331, p-value = 0.000, r = -0.087), and Intermediate and Advanced (-1,459.402, z = -3.685, p-value = 0.001, r = -0.024) ELs. In these cases, Advanced ELs accessed Sticky Notes significantly more

than did Beginner and Intermediate ELs, and Intermediate ELs accessed the tool significantly more than did Beginners. Yet, these differences were not meaningful as demonstrated by small effect sizes.

Table 30. Pairwise comparison results in the writing domain

		U	*z*	*p*-value	*r*
Beginner vs. Intermediate	Color Overlay	-169.709	-2.593	0.029	-0.024
	Color Contrast	-253.212	-3.523	0.001	-0.032
	Highlighter	-604.551	-4.273	0.000	-0.023
	Sticky Notes	-730.055	-5.785	0.000	-0.036
Beginner vs. Advanced	Sticky Notes	-2,189.457	-5.331	0.000	-0.087
Intermediate vs. Advanced	Sticky Notes	-1,459.402	-3.685	0.001	-0.024

In conclusion, Intermediate and Advanced ELs generally made high use of all tools in all domains, whereas Beginner ELs displayed the lowest tool activation rates. Despite variations in the percentage of a PL group that accessed any given tool, the median access rates were similar across all PLs. Significant differences were observed in the rates at which different PL groups accessed some tools, yet the effect sizes were small, suggesting that the significant differences could be attributed to relatively large sample sizes.

4.3.2. Use of Tools across Different Proficiency Levels by Grade-Level Clusters

Next, the activation of tools across different PLs in each grade-level cluster is summarized. In the listening domain (Table 31), considerable differences in tool use existed among PLs in the Grades 4-5 and Grades 6-8 grade-level clusters. For example, in the Grades 4-5 group, 3% more Intermediate than Beginner ELs activated the Magnifier. Highlighter activation was, similarly, 2% higher among Intermediate than Beginner ELs. Additionally, within the Grades 6-8 group, 2% more Intermediate than Advanced ELs activated the Line Guide. Activation of the two Color tools was also somewhat different between Beginner ELs and the Intermediate and Advanced groups in the Grades 4-5 and Grades 6-8 grade-level clusters.

About 1% more Intermediate and Advanced students accessed the Color tools than did Beginners. Tool activation within each PL in Grades 9-12 was similar with respect to the activation of the two Color tools, the two Help tools, and the Highlighter, with less than 1% difference between the PL groups. The largest difference among PLs within this grade-level cluster was in activation of the Line Guide. In comparison to Advanced students, 3% more Beginner and 2% more Intermediate students activated the Line Guide. Activation of the Magnifier and Help (General) tools was also somewhat different between Beginner and Advanced ELs in this grade-level cluster. While 1.5% more Beginner ELs accessed the Help (General) tool, 1.5% more Advanced ELs activated the Magnifier tool. On the other hand, universal tool activation rates did not vary much across the three PLs within the Grade 1 and Grades 2-3 grade-level clusters. In general, the percentage of Beginner, Intermediate, and Advanced students using the tools in these grade-level clusters was very similar. Activation of the Line Guide was an exception within the Grade 1 group, as 2% more Beginner than Advanced ELs activated the tool.

In the reading domain (Table 32), ELs in different PLs in the Grades 4-5, Grades 6-8, and Grades 9-12 grade-level clusters showed comparatively more variation in their universal tool activation rates. Specifically, the difference between Beginner and Advanced students was relatively large with respect to activation of the Highlighter and Line Guide tools. Regarding Highlighter access, the difference was about 7% in Grades 9-12, 6% in Grades 6-8, and 3% in Grades 4-5, with a higher percentage of Advanced ELs making use of the tool than Beginners. Line Guide activation among Advanced ELs was 4% more than that of Beginner ELs in Grades 4-5, 3.7% more in Grades 6-8, and 2% more in Grades 9-12. On the other hand, while 2% more Advanced than Beginner ELs Grades 9-12 activated the Magnifier, within the Grades 4-5 and

Grades 6-8 groups, 2% more Intermediate than Beginner ELs activated the tool. Similarly, activation of two Color tools was slightly higher among Intermediate ELs than Beginners in Grades 4-5 and Grades 6-8, but slightly higher among Advanced ELs than Beginners in Grades 9-12. On the other hand, activation of the tool was similar among Grade 1 and Grades 2-3 students in the reading domain. The largest differences occurred in activation of the Line Guide, Highlighter, and Magnifier tool between Beginner and Advanced ELs in these grades. Specifically, with respect to Line Guide, the difference between Beginner and Advanced EL activation of the tool was 1.8% in Grade 1 and 1.5% in Grades 2-3.

In the speaking domain (Table 33), tool activation across PLs was generally more similar in each grade. There were some exceptions. For example, in the Grades 2-3 grade-level cluster, the largest difference was observed between Beginner and Advanced students in activation of the Highlighter (3.8%), Magnifier (2.6%), and Color Contrast (2%) tools, with more Advanced ELs activating the tool. The difference in Line Guide activation in Grades 4-5 and Grades 9-12 was also somewhat significant. In Grades 4-5, 2.3% more Advanced than Beginner ELs accessed the tool. On the other hand, in Grades 9-12, 1.5% more Beginner and Intermediate ELs than Advanced ELs accessed the Line Guide.

In the writing domain (Table 34), tool activation varied more, particularly between Grades 6-8 Beginner and Advanced ELs. Specifically, the differences were 2.3% for Color Overlay, 3.1% for Line Guide, 4.5% for Highlighter, about 2% for Magnifier, and 8% for Sticky Notes, with a higher percentage of Advanced ELs activating the tools. Additionally, comparatively large differences in Highlighter and Sticky Notes activation rates were observed within the Grades 4-5 and Grades 9-12 groups. In both grade-level clusters, 5% more Advanced than Beginner ELs accessed Sticky Notes. With respect to the Highlighter, while 5% more

Advanced than Beginner ELs in Grades 9-12 accessed the tool, the difference was larger between Intermediate and Beginner ELs in Grades 4-5, with 2% more Intermediate ELs activating the tool.

In conclusion, more variation was observed among proficiency levels—particularly between the Beginner and Intermediate levels—in the Grades 4-5 listening domain (a significantly higher percentage of Intermediate ELs activated the tools). The difference in Line Guide activation within the Grade 1, Grades 6-8, and Grades 9-12 groups was comparatively large, as 2% more Intermediate than Beginner ELs accessed the tool in each of these grade-level clusters. In the reading domain, the differences in tool activation across PLs were more apparent among the Grades 4-5, Grade 6-8, and Grade 9-12 groups. Specifically, Beginner and Advanced ELs differed in activation of the Highlighter, Line Guide, and Magnifier in these grade-level clusters, with more Advanced ELs making use of these tools. Tool access was generally similar across PLs in the speaking domain. The Grades 2-3 group was an exception, especially in accessing the Color Contrast, Highlighter, and Magnifier tools, as 2% or more of Advanced ELs activated these tools. In the writing domain, Beginner and Advanced learners, specifically in Grades 6-8, showed variation in their tool activation. More Advanced students made use of the tools in general, and the difference was particularly large for the Sticky Notes and Highlighter tools. The difference between Beginner and Advanced learners was also large in Sticky Notes activation in the Grades 4-5 and Grades 9-12 grade-level clusters, with notably more Advanced students using these tools.

Table 31. Universal tool use by PL within grade-level clusters in the listening domain

Grade levels	PL	Color Overlay				Color Contrast				Help (General)				Help (Tools)			
		Use %	$\bar{X}$	*Med*	*sd*	Use %	$\bar{X}$	*Med*	*sd*	Use %	$\bar{X}$	*Med*	*sd*	Use %	$\bar{X}$	*Med*	*sd*
1	**Beginner** (*n* = 25,357)	0.7	1.79	1	1.173	0.7	2.18	1	1.96	1.8	1.22	1	0.771	0.8	1.2	1	0.56
	Intermediate (*n* = 29,554)	0.8	1.98	1	1.477	0.8	1.94	1	1.433	2	1.18	1	0.495	0.7	1.17	1	0.604
	Advanced (*n* = 120,915)	0.6	1.72	1	1.274	0.7	1.9	1	1.449	2.1	1.12	1	0.403	0.8	1.08	1	0.39
2-3	**Beginner** (*n* = 71,647)	1.3	1.79	1	1.237	1.4	1.9	1	1.408	2.9	1.15	1	0.48	1.3	1.12	1	0.428
	Intermediate (*n* = 85,586)	1.5	1.81	1	1.276	1.5	1.91	1	1.369	2.8	1.11	1	0.416	1.3	1.1	1	0.394
	Advanced (*n* = 207,883)	1.6	1.74	1	1.128	1.6	1.79	1	1.227	2.4	1.09	1	0.333	1.4	1.09	1	0.392
4-5	**Beginner** (*n* = 16,718)	1.7	1.74	1	1.187	1.6	2	2	1.402	3.6	1.15	1	0.459	1.9	1.16	1	0.612
	Intermediate (*n* = 34,814)	2.7	1.9	1	1.379	2.7	1.98	2	1.329	3.6	1.12	1	0.47	2.1	1.15	1	0.579
	Advanced (*n* = 165,405)	2.8	1.69	1	1.107	2.7	1.79	1	1.247	2.7	1.09	1	0.357	1.7	1.11	1	0.421
6-8	**Beginner** (*n* = 35,479)	4	1.77	1	1.26	3.5	1.95	1	1.477	4.9	1.22	1	0.689	3.3	1.18	1	0.642
	Intermediate (*n* = 87, 348)	5.3	1.73	1	1.137	4.8	1.86	1	1.326	4	1.11	1	0.388	3.4	1.12	1	0.473
	Advanced (*n* = 119,543)	5.1	1.69	1	1.087	4.8	1.8	1	1.176	3.1	1.1	1	0.38	2.6	1.1	1	0.391
9-12	**Beginner** (*n* = 78,258)	2.2	1.54	1	1.010	1.6	1.54	1	0.95	3.6	1.18	1	0.763	2.3	1.14	1	0.516
	Intermediate (*n* = 111,741)	2.9	1.51	1	0.912	2.3	1.55	1	0.927	2.8	1.08	1	0.333	2.3	1.06	1	0.296
	Advanced (*n* = 71,308)	2.9	1.51	1	0.860	2.4	1.57	1	1.037	2.1	1.08	1	0.332	1.9	1.05	1	0.27

(Table continues)

Table 31 Universal tool use by PL within grade-level clusters in the listening domain (continued)

Grade levels	PL	Line Guide				Highlighter				Magnifier			
		Use %	$\bar{X}$	*Med*	*sd*	Use %	$\bar{X}$	*Med*	*sd*	Use %	$\bar{X}$	*Med.*	*sd*
1	**Beginner** (*n* = 25,357)	6	1.97	1	3.757	2.4	16.75	7	26.321	6.1	32.6	2	40.06
	Intermediate (*n* = 29,554)	5	1.74	1	1.84	2.6	14.42	6	29.695	5.8	36.96	2	46.401
	Advanced (*n* = 120,915)	3.9	1.63	1	1.591	2.2	11.08	5	20.682	5.4	45.03	4	48.505
2-3	**Beginner** (*n* = 71,647)	6.5	1.99	1	2.384	4.3	12.88	6	22.704	8.7	22.08	2	37.088
	Intermediate (*n* = 85,586)	6.2	1.87	1	1.672	4.5	10.33	5	16.212	8.4	25.97	2	41.911
	Advanced (*n* = 207,883)	6.2	1.92	1	1.735	4.4	9.48	5	15.138	8.6	27.12	2	42.695
4-5	**Beginner** (*n* = 16,718)	8.5	2.15	1	2.675	4.9	13.33	6	23.189	10.7	15.37	2	31.227
	Intermediate (*n* = 34,814)	9.6	1.97	1	1.657	6.8	11.47	6	20.108	13.7	15.24	2	33.303
	Advanced (*n* = 165,405)	8.9	1.94	1	1.712	5.6	9.77	5	15.645	12	19.09	2	37.081
6-8	**Beginner** (*n* = 35,479)	14	2.23	1	3.168	7.1	15.74	7	32.968	12.7	10.68	2	26.764
	Intermediate (*n* = 87, 348)	12.9	1.98	1	1.787	7.5	15.14	7	28.262	13.6	11.6	2	29.245
	Advanced (*n* = 119,543)	11.8	1.97	1	1.928	6.6	15.09	6	30.87	12.8	15.49	2	33.894
9-12	**Beginner** (*n* = 78,258)	12.7	2.06	1	8.943	3.6	12.42	5	21.34	5.9	14.94	2	31.156
	Intermediate (*n* = 111,741)	10.7	1.8	1	1.42	3.5	13.81	6	26.177	7	15.95	2	34.382
	Advanced (*n* = 71,308)	9.8	1.8	1	1.537	3.6	13.94	6	28.006	7.3	18.94	2	36.935

Table 32. Universal tool use by PL within grade-level clusters in the reading domain

Grade levels	PL	Color Overlay				Color Contrast				Help (General)				Help (Tools)			
		Use %	$\bar{X}$	*Med*	*sd*	Use %	$\bar{X}$	*Med*	*sd*	Use %	$\bar{X}$	*Med*	*sd*	Use %	$\bar{X}$	*Med*	*sd*
1	**Beginner** (*n* = 79,415)	0.6	1.63	1	1.219	0.6	1.72	1	1.101	0.9	1.17	1	0.528	0.4	1.1	1	0.358
	Intermediate (*n* = 50,222)	0.6	1.6	1	1.09	0.6	1.68	1	1.193	1	1.12	1	0.497	0.4	1.11	1	0.318
	Advanced (*n* = 46,126)	0.7	1.6	1	0.96	0.7	1.6	1	0.977	0.7	1.14	1	0.456	0.3	1.1	1	0.34
2-3	**Beginner.** (*n* = 170,858)	1.5	1.7	1	1.174	1.6	1.84	1	1.276	1.2	1.14	1	0.524	0.6	1.18	1	0.606
	Intermediate (*n* = 91,879)	1.5	1.68	1	1.274	1.5	1.83	1	1.387	0.6	1.15	1	0.545	0.4	1.13	1	0.405
	Advanced (n= 102,202)	1.5	1.65	1	1.145	1.6	1.77	1	1.347	0.4	1.12	1	0.423	0.3	1.22	1	0.621
4-5	**Beginner** (*n* = 83,622)	2.9	1.77	1	1.269	3	1.91	1	1.463	1.4	1.16	1	0.487	0.9	1.18	1	0.563
	Intermediate (*n* = 71,865)	3.5	1.73	1	1.168	3.6	1.91	1	1.366	1.1	1.13	1	0.441	0.9	1.26	1	0.703
	Advanced (*n* = 61,344)	3.3	1.71	1	1.122	3.5	1.9	1	1.306	0.8	1.11	1	0.422	0.7	1.23	1	0.578
6-8	**Beginner** (*n* = 142,075)	3.1	1.58	1	1	3.2	1.73	1	1.17	1	1.12	1	0.493	0.7	1.2	1	0.651
	Intermediate (*n* = 60,246)	3.8	1.64	1	1.063	4	1.77	1	1.294	0.7	1.1	1	0.357	0.7	1.16	1	0.487
	Advanced (*n* = 40,920)	3.6	1.62	1	1.038	3.7	1.77	1	1.262	0.7	1.11	1	0.373	0.7	1.25	1	0.589
9-12	**Beginner** (*n* = 140,402)	1.2	1.48	1	0.887	1	1.52	1	0.932	0.4	1.13	1	0.391	0.3	1.13	1	0.471
	Intermediate (*n* = 55,916)	1.7	1.51	1	0.985	1.5	1.64	1	1.092	0.4	1.04	1	0.195	0.3	1.1	1	0.379
	Advanced (*n* = 64,302)	1.8	1.48	1	0.882	1.7	1.61	1	1.034	0.4	1.11	1	0.45	0.3	1.11	1	0.473

(Table continues)

Table 32 Universal tool use by PL within grade-level clusters in the reading domain (continued)

Grade levels	PL	Line Guide				Highlighter				Magnifier			
		Use %	$\bar{X}$	*Med*	*sd*	Use %	$\bar{X}$	*Med*	*sd*	Use %	$\bar{X}$	*Med*	*sd*
1	**Beginner** (*n* = 79,415)	4	1.84	1	2.454	3	14.64	7	25.728	5.2	46.6	4	52.019
	Intermediate (*n* = 50,222)	4.1	2.25	1	3.521	3.3	12.38	6	18.767	5.5	48.46	4	54.478
	Advanced (*n* = 46,126)	4.8	2.82	1	3.735	4.3	14.07	7	22.37	6	49.77	6	54.583
2-3	**Beginner** (*n* = 170,858)	7.1	2.86	2	3.43	8.2	15.11	8	24.47	9.3	25.27	2	43.486
	Intermediate (*n* = 91,879)	7.8	3.47	2	4.087	8.8	14.63	7	23.871	9	29.06	2	46.95
	Advanced (*n* = 102,202)	8.6	3.9	2	4.654	9.2	15	8	23.822	8.9	32.19	2	48.653
4-5	**Beginner** (*n* = 83,622)	11.2	2.99	2	3.25	14.9	16.95	9	26.501	15.2	15.32	2	34.289
	Intermediate (*n* = 71,865)	14.3	3.42	2	3.845	18.3	16.79	9	25.318	17.3	16.87	2	36.435
	Advanced (*n* = 61,344)	15.2	3.68	2	4.278	18.3	17.49	9	28.7	16.2	19.06	2	38.802
6-8	**Beginner** (*n* = 142,075)	7.7	2.18	1	2.248	11.7	18.02	9	29.803	9.9	16.13	2	35.166
	Intermediate (*n* = 60,246)	9.6	2.46	2	2.682	16.1	18.97	9	31.689	11.2	18.48	2	39.217
	Advanced (*n* = 40,920)	10	2.67	2	5.776	17.4	20.55	9	34.464	10.8	24.45	2	43.98
9-12	**Beginner** (*n* = 140,402)	4	1.92	1	2.042	5	18	8	31.311	4.4	24.89	2	42.018
	Intermediate (*n* = 55,916)	4.9	2.12	1	2.31	9.4	20.23	10	34.018	6	23.91	2	43.392
	Advanced (*n* = 64,302)	6.1	2.36	1	3.19	12.1	23.15	10	38.98	6.6	23.06	2	42.468

Table 33. Universal tool use by PL within grade-level clusters in the speaking domain

Grade levels	PL	Color Overlay				Color Contrast				Help (General)				Help (Tools)			
		Use %	$\bar{X}$	*Med.*	*sd*	Use %	$\bar{X}$	*Med.*	*sd*	Use %	$\bar{X}$	*Med.*	*sd*	Use %	$\bar{X}$	*Med.*	*sd*
1	**Beginner** (*n* = 72,764)	0.5	1.64	1	1.031	0.5	1.77	1	1.266	0.7	1.13	1.0	0.563	0.5	1.1	1.00	0.464
	Intermediate (*n* = 99,657)	0.5	1.53	1	0.993	0.5	1.66	1	1.175	0.6	1.08	1.0	0.285	0.4	1.11	1.00	0.47
	Advanced (*n* = 330)	0.0	NA	NA	NA	0.3	1	1	NA	0	NA	NA	NA	0	NA	NA	NA
2-3	**Beginner** (*n* = 182,302)	0.9	1.57	1	1.108	1	1.63	1	1.18	0.7	1.13	1.0	0.501	0.5	1.17	1.00	0.652
	Intermediate (*n* = 178,352)	1.0	1.52	1	1.016	1.1	1.63	1	1.31	0.5	1.09	1.0	0.374	0.4	1.11	1.00	0.379
	Advanced (*n* = 468)	1.9	1.89	1	1.833	3	1.64	1	1.082	0.4	1.5	1.5	0.707	0.2	0	1.00	NA
4-5	**Beginner** (*n* = 70,198)	1.9	1.59	1	1.174	2.1	1.69	1	1.26	1.3	1.12	1.0	0.441	1	1.2	1.00	0.645
	Intermediate (*n* = 142,774)	1.9	1.51	1	0.896	2.2	1.59	1	1.07	0.8	1.1	1.0	0.357	0.7	1.16	1.00	0.508
	Advanced (*n* = 1,360)	1.8	1.42	1	0.776	2.4	1.67	1	1.021	0.4	1	1.0	0	0.9	1.33	1.00	0.651
6-8	**Beginner** (*n* = 104,100)	2.9	1.53	1	0.954	2.8	1.59	1	1.054	2	1.11	1.0	0.406	1.7	1.18	1.00	0.592
	Intermediate (*n* = 133,425)	2.6	1.49	1	.901	2.7	1.51	1	0.881	1.6	1.08	1.0	0.326	1.2	1.15	1.00	0.493
	Advanced (*n* = 459)	2.2	1,1	1	.316	1.7	1.75	1	1.035	0.7	1	1.0	0.000	0.7	1	1.00	0
9-12	**Beginner** (*n* = 162,999)	1.2	1.46	1	.846	1	1.47	1	0.967	0.9	1.09	1.0	0.377	0.8	1.12	1.00	0.425
	Intermediate (*n* = 88,541)	1.3	1.39	1	.786	1.1	1.46	1	1.124	0.8	1.1	1.0	0.379	0.7	1.13	1.00	0.451
	Advanced (*n* = 73)	0.0	NA	NA	NA	1.4	1.00	1	NA	1.4	1	1.0	0.000	0	NA	NA	NA

(Table continues)

Table 33. Universal tool use by PL within grade-level clusters in the speaking domain (continued)

Grade level	PL	Line Guide				Highlighter				Magnifier			
		Use %	$\bar{X}$	*Med*	*sd*	*Use %*	$\bar{X}$	*Med*	*sd*	*Use %*	$\bar{X}$	*Med*	*sd*
1	**Beginner** (*n* = 72,764)	3.0	1.67	1.00	2.636	1.9	12.06	6.0	19.225	4	68.24	52	66.086
	Intermediate (*n* = 99,657)	2.8	1.63	1.00	1.623	2.1	9.68	5.0	13.466	4	84.9	112	75.879
	Advanced (*n* = 330)	3.3	1.55	1.00	0.688	1.8	5.17	4.5	4.167	3.9	87.85	154	83.188
2-3	**Beginner** (*n* = 182,302)	3.6	1.85	1.00	1.746	3.3	9.64	5.0	16.77	4.9	50.09	2.0	63.976
	Intermediate (*n* = 178,352)	4.0	1.93	1.00	2.016	3.8	8.33	5.0	13.908	5.3	56.53	2.0	69.445
	Advanced (*n* = 468)	4.1	2.37	2.00	2.385	7.1	10.42	5.0	16.941	7.5	51.4	2.0	69.783
4-5	**Beginner** (*n* = 70,198)	5.6	2.07	1.00	2.102	5.4	10.56	5.0	20.238	7.2	30.61	2.0	50.611
	Intermediate (*n* = 142,774)	5.8	1.99	1.00	2.801	5.1	8.73	4.0	17.608	7	34.89	2.0	55.681
	Advanced (*n* = 1,360)	7.9	2.06	1.00	2.437	6	6.27	3.0	7.187	7.6	21.25	1.0	45.9
6-8	**Beginner** (*n* = 104,100)	6.2	1.94	1.00	1.785	5.9	15.32	6.0	35.452	6.5	29.62	2.0	55.406
	Intermediate (*n* = 133,425)	6.1	1.97	1.00	1.943	5.4	13.04	6.0	26.721	6.3	41.19	2.0	65.653
	Advanced (*n* = 459)	5.9	1.37	1.00	0.884	5.9	9.11	6.0	21.484	6.3	52.45	2.0	71.408
9-12	**Beginner** (*n* = 162,999)	3.9	1.75	1.00	1.585	2.6	17.13	6.0	39.421	3.3	40.49	2.0	55.822
	Intermediate (*n* = 88,541)	3.9	1.75	1.00	1.541	2.8	17.22	6.0	47.93	3.4	43.83	2.0	60.001
	Advanced (*n* = 73)	1.4	2.00	2.00	NA	2.7	7.5	7.5	9.192	2.7	68.5	68.5	86.974

Table 34. Universal tool use by PL within grade-level clusters in the writing domain

Grade level	PL	Color Overlay				Color Contrast				Help (General)				Help (Tools)			
		Use %	$\bar{X}$	*Med.*	*sd*	Use %	$\bar{X}$	*Med.*	*sd*	Use %	$\bar{X}$	*Med.*	*sd*	Use %	$\bar{X}$	*Med.*	*sd*
4-5	**Beginner** (*n* = 21,972)	1.5	1.59	1	1.162	1.6	1.79	1	1.451	1.6	1.11	1	0.394	0.8	1.15	1	0.435
	Intermediate (*n* = 138,251)	2.1	1.65	1	1.159	2.5	1.82	1	1.3	0.8	2.09	1	0.34	0.7	1.18	1	0.551
	Advanced (*n* = 1,735)	2.2	1.58	1	1.030	2.2	1.77	1	1.347	1.1	1.16	1	0.501	0.6	1.1	1	0.316
6-8	**Beginner** (*n* = 71,648)	1.9	1.50	1	1.058	2.2	1.61	1	1.143	1.2	1.11	1	0.368	0.8	1.16	1	0.548
	Intermediate (*n* = 167,368)	2.6	1.50	1	.877	3	1.65	1	1.068	0.5	1.09	1	0.368	0.5	1.19	1	0.655
	Advanced (*n* = 595)	4.2	1.56	1	.821	3.4	1.2	1	0.41	0.5	1.00	1	0	1.3	1.13	1	0.354
9-12	**Beginner** (*n* = 73,397)	0.7	1.39	1	.768	0.6	1.45	1	0.854	0.8	1.11	1	0.545	0.5	1.08	1	0.382
	Intermediate (*n* = 181,233)	1.0	1.38	1	.753	0.9	1.48	1	0.895	0.4	1.08	1	0.342	0.3	1.1	1	0.379
	Advanced (*n* = 1,743)	1.3	1.23	1	.612	1.3	1.59	1	0.908	0.3	1.00	1	0	0.6	1.18	1	0.405

Grade Level	PL	Line Guide				Highlighter				Magnifier				Sticky Notes			
		Use %	$\bar{X}$	*Med.*	*sd*	Use %	$\bar{X}$	*Med.*	*sd*	Use %	$\bar{X}$	*Med.*	*sd*	Use %	$\bar{X}$	*Med.*	*sd*
4-5	**Beginner** (*n* = 21,972)	4.1	1.62	1	1.261	5.3	12.19	5	26.021	6.1	15.45	2	28.715	2.9	2.08	1.0	2.48
	Intermediate (*n* = 138,251)	4.9	1.65	1	1.397	7.3	10.51	6	18.327	7.2	16	2	27.998	6	2.04	1.0	2.069
	Advanced (*n* = 1,735)	6.2	1.72	1	1.433	7	11.35	6	14.646	7.7	17.07	2	26.163	7.9	2.18	1.0	2.043
6-8	**Beginner** (*n* = 71,648)	4	1.66	1	2.46	4.7	13.29	6	29.718	4.5	16.04	2	25.129	2.9	1.98	1.0	1.386
	Intermediate (*n* = 167,368)	4.2	1.65	1	1.471	6.3	13.54	7	23.836	5.6	16.05	2	24.864	5.6	2.08	1.0	2.666
	Advanced (*n* = 595)	7.1	1.5	1	0.804	9.2	11.93	7	14.21	7.4	12.25	2	23.124	10.8	2.19	1.5	1.885
9-12	**Beginner** (*n* = 73,397)	3.3	1.63	1	1.4	2	12.02	5	25.785	2.6	21.06	2	31.486	1	1.69	1.0	2.734
	Intermediate (*n* = 181,233)	3.3	1.58	1	1.185	4.2	12.16	5	30.834	3.4	16.14	2	23.461	3	1.92	1.0	1.865

4.4. Findings from Research Question 4

Research Question 4 asks: To what extent do different item types and features affect ELs' use of the accessibility features embedded in an ELP assessment?

4.4.1. Use of Tools at the Item Level

This section summarizes the preliminary findings for item-level tool activation. Tables 35-38 are based on the percentage of students activating each tool for each of the items presented

in a domain. The tables include the maximum and minimum percentage of students accessing a given tool across all items of the domain (denoted by *max* and *min*), the average percentage of students activating a given tool across all items of the domain (denoted by *mean*), and the variation of the percentage of students activating a given tool across all items of the domain (denoted by *sd)*. In each domain, one to five items that triggered more tool use than other items are noted for potential further exploration. These items are the ones for which a high percentage of students using some combination of the Line Guide, Highlighter, and Magnifier tools while responding to the items. Certain characteristics of items (e.g., item type, difficulty, position of the item in the test) might have caused increased use of the tools.

Overall, compared to the findings relevant to research questions 1 through 3, tool activation at the item level was low. This finding is expected because the analysis is now evaluating the number of activations of each tool for individual items rather than aggregate use of those tools throughout the domain. While a student might activate a tool within a given domain, the student does not necessarily have that tool active while responding to all items in the domain. Furthermore, the research team anticipated that not all of the items would trigger the same amount of tool use, a hypothesis that was supported by the wide range of the percentage of students who activated the tools for each item (see the *max* and *min* values in Tables 35-38). Findings are presented for each domain, including details of tool activation for each grade-level cluster.

Item level tool activation was generally low in the listening domain (Table 35). Except for the Magnifier, the average percentage of students accessing a tool while responding to an item is less 1%, and there is not much variation across items. Among Grade 1 students, about 2% consistently activated the Magnifier as they moved from item to item. The use of other tools is

fairly limited. On average, per item, only 0.07% of ELs activated the Highlighter, and only 0.1% activated the Line Guide. The least accessed tools per item were the Color and Help tools. On average, per item, only about 0.02% of students activated any of these four tools. Due to consistent, low activation of tools across items, there might be no further need to explore specific items for this grade-level cluster.

Grades 2-3 students displayed a tool activation pattern very similar to that of students in Grade 1. About 2% of the Grades 2-3 students used the Magnifier while responding to items. For one specific item, 2.7% used this tool. Similarly, the mean percentage of students accessing the Highlighter (0.1%) and Line Guide (0.2%) tools was slightly higher in the Grades 2-3 group than it was in the Grade 1 group. There were few items, such as item *12815, 11546, 12953, 12707, 12971,* for which a relatively high percentage of students activated a combination of the Line Guide, Highlighter, and Magnifier tools.

In general, tool activation at the item level among Grades 4-5 students was consistent with the above findings. Again, about 2% of the students, on average, activated the Magnifier for any individual item. Compared to other grades, the Grades 4-5 group showed more deviation across items for activation of the Magnifier (*max* = 3.1%, *min* = 1.7%, and *sd* = 0.3). Additionally, the average percentage of students accessing tools per item was slightly higher, especially for the Highlighter (0.2%, *max* = 0.6%) and Line Guide (0.3%, *max* = 0.8%). Particular use of these tools was consistent with domain-specific findings for Research Question 1. Items *12585, 12598,* and *13032* might warrant further analysis, as the percentage of students activating the Line Guide, Highlighter, and Magnifier was higher for these items than for any of the other items in the domain.

In the Grades 6-8 group, the average percentage of students accessing the Magnifier was slightly lower (*mean* = 1.9%, *max* = 2.6%, *min* = 1.4%). On the other hand, although still limited, the mean percentage of students activating the Highlighter (0.3%) and Line Guide (0.4%) per item was slightly higher than in the Grades 4-5 group. Item *13044* might be of particular interest as many students activated all three tools—the Line Guide, Highlighter, and Magnifier—for this item.

Finally, in the Grades 9-12 group the mean percentage of students activating the Magnifier (1.3%), Line Guide (0.3%), and Highlighter (0.1%) for an item was lower than in the Grades 4-5 and Grades 6-8 groups. Specifically, activation of the Magnifier for each item was lower in Grades 9-12 than in any other grade-level cluster. Also, the items for which a notably high percentage of students activated the Highlighter, Line Guide, and Magnifier tools varied widely, with the exception of item *12720*, which might therefore warrant further investigation.

In conclusion, although item level tool access was low in the listening section, certain items attracted more students than other items, especially in the use of the Highlighter, Line Guide, and Magnifier. In general, students were using a combination of these tools for those items. Thus, a more detailed analysis of these specific items is suggested.

Table 35. Universal tool use at the item level in the listening domain

Grade levels		Color Overlay %	Color Contrast %	Help (General) %	Help (Tools) %	Highlighter %	Line Guide %	Magnifier %
1	*Max*	0.054	0.055	0.155	0.052	0.251	0.340	2.491
	Min	0.000	0.004	0.004	0.000	0.019	0.053	2.206
	Mean	0.019	0.021	0.044	0.018	0.065	0.141	2.377
	SD	0.011	0.011	0.029	0.010	0.041	0.065	0.080
2-3	*Max*	0.135	0.138	0.680	0.404	0.359	0.612	2.694
	Min	0.012	0.019	0.007	0.000	0.045	0.077	2.046
	Mean	0.046	0.047	0.068	0.033	0.139	0.209	2.302
	SD	0.025	0.025	0.091	0.053	0.071	0.106	0.167

Grade levels		Color Overlay %	Color Contrast %	Help (General) %	Help (Tools) %	Highlighter %	Line Guide %	Magnifier %
4-5	*Max*	0.276	0.254	0.207	0.116	0.573	0.798	3.091
	Min	0.008	0.000	0.010	0.000	0.036	0.114	1.709
	Mean	0.079	0.081	0.086	0.051	0.193	0.308	2.261
	SD	0.058	0.057	0.052	0.028	0.119	0.160	0.324
6-8	*Max*	0.352	0.385	0.352	0.301	0.552	0.798	2.563
	Min	0.044	0.020	0.022	0.005	0.064	0.122	1.432
	Mean	0.163	0.157	0.102	0.080	0.266	0.420	1.879
	SD	0.072	0.077	0.079	0.058	0.105	0.154	0.277
9-12	*Max*	0.148	0.120	0.281	0.203	0.204	0.742	1.603
	Min	0.020	0.015	0.009	0.005	0.034	0.117	0.974
	Mean	0.075	0.059	0.065	0.051	0.117	0.294	1.263
	SD	0.032	0.027	0.057	0.041	0.045	0.132	0.183

In the reading domain, tool activation at the item level is generally low (Table 36), as in the listening domain. Like in the domain-level findings, students accessed the Magnifier more often than other tools at the item level. There is also variation in tool use among items. Specific items initiated more use of the Line Guide, Highlighter, and Magnifier than other items. Variation also occurred across grade-level clusters. On average, more students activated the Line Guide per item in the Grades 4-5 group. Similarly, the average frequency of Magnifier access per item was lowest in the Grades 9-12 group. Tool activation is broken down at the item level for each grade-level cluster along with some specific items flagged for potential further investigation.

In Grade 1, the average frequency at which students accessed a given tool for an item was lower than in the other grade-level clusters, and tool use per item, overall, was very limited. One exception was activation of the Magnifier. On average, 2.4% of the Grade 1 students used the Magnifier while responding to an item. With respect to the Highlighter and Line Guide, an average of about 0.1% of Grade 1 students activated the tool for an item. For a few items, about 0.3% of the students activated the Highlighter and about 0.4% used the Line Guide. As seen by the minimum percentage values, for some items there is no tool activation, especially of the

Color and Help tools. Closer inspection of items found that concurrent activation of the Line Guide, Highlighter, and Magnifier was particularly high for three of the items (i.e., items *14620, 14621*, and *13219*).

Compared to tool activation among Grade 1 students, tool access at the item level among Grades 2-3 students was slightly higher. Magnifier use was more consistent across items, with about 2.1% of the students activating that tool per item. For few items, about 3% of students accessed this tool. Additionally, about 0.4% of the students used the Highlighter and Line Guide when responding to an item. For two or three items the maximum percentage of students activating these tools was about 1.2%. As in Grade 1, a particular item stood out for high tool activation. A notably large percentage of students activated both the Line Guide and the Highlighter while responding to item *13367*. Items *13932, 13327,* and *13366* may warrant further exploration as these items triggered more use of a combination of the Line Guide, Highlighter, and Magnifier.

On average, a higher percentage of Grades 4-5 students, relative to students in other grade-level clusters, accessed one or more tools while answering any given item. On average. 2.7% of the students used the Magnifier per item. The maximum percentage of the students turning on this tool increased to about 4.5% for two of the items. The mean percentage of students using the Highlighter and Line Guide for an item was about 0.8%. For five of the items, about 2% of students activated the Highlighter. For the Line Guide, in addition to four items that triggered use by 2% of the students, one item triggered 3% of students to activate the tool. Variation with respect to activation of the Highlighter (sd = 0.534), Line Guide (sd = 0.67), and Magnifier (sd = 0.57), per item, was largest within this grade-level cluster. Additionally, items

13490, 13482, 13926, 13928, and *13450* triggered for the most activation by students of a combination of the Line Guide, Highlighter, and Magnifier.

Activation of the tools at the item level in Grades 6-8 was similar to that in Grades 4-5, except for activation of the Line Guide and Magnifier. Magnifier access per item, on average, was lower among the Grades 6-8 students than among the Grades 4-5 students. On average, 2% of the Grades 6-8 students accessed the Magnifier while responding to an item. For few items, 2.7% of the students used this tool. Also, 0.4% of students, on average, activated the Line Guide for an item. For three items, 1.1% of students used this tool. Items *14617, 13614, 13631, 14492,* and *13659* might be further explored as more students accessed a combination of the Highlighter, Line Guide, and Magnifier for these items than for others.

Finally, among Grades 9-12 students, on average, fewer students activated a given tool compared to students in Grades 4-5 or Grades 6-8. The average activation of the Magnifier per item was lowest among Grades 9-12 students at 1.3%. Only for one item did 2% of the students use the Magnifier. The mean percentage of students activating the Highlighter for an item was 0.6%. However, there was one item for which about 3% of the students activated this tool. For another item, about 2% of the students used the tool. For the Line Guide, about 0.2% of students, on average, activated the tool when responding to an item. Items *13719, 13720,* and *13970* can be further explored as a high percentage of students used all three tools: Highlighter, Line Guide, and Magnifier.

In conclusion, average item level tool activation was limited in the reading domain. In contrast to the listening domain, more students accessed the Line Guide, Highlighter, and Magnifier on some items, and that finding requires further exploration.

Table 36. Universal tool use per item in reading domain

Grade levels		Color Overlay %	Color Contrast %	Help (General) %	Help (Tools) %	Highlighter %	Line Guide %	Magnifier %
1	*Max*	0.028	0.027	0.039	0.016	0.271	0.393	2.700
	Min	0.000	0.000	0.000	0.000	0.022	0.035	2.193
	Mean	0.008	0.008	0.008	0.004	0.103	0.132	2.410
	SD	0.005	0.006	0.007	0.004	0.064	0.078	0.175
2-3	*Max*	0.087	0.100	0.185	0.068	1.172	1.326	2.959
	Min	0.001	0.005	0.001	0.000	0.069	0.051	2.059
	Mean	0.029	0.032	0.014	0.008	0.391	0.420	2.407
	SD	0.016	0.018	0.022	0.009	0.239	0.294	0.225
4-5	*Max*	0.255	0.298	0.106	0.063	2.381	3.004	4.488
	Min	0.018	0.012	0.003	0.000	0.215	0.081	1.795
	Mean	0.076	0.086	0.027	0.018	0.841	0.757	2.696
	SD	0.053	0.061	0.020	0.014	0.534	0.607	0.579
6-8	*Max*	0.239	0.273	0.077	0.064	2.481	1.117	2.767
	Min	0.021	0.024	0.002	0.004	0.153	0.081	1.316
	Mean	0.083	0.093	0.020	0.017	0.842	0.389	1.965
	SD	0.051	0.058	0.014	0.011	0.556	0.280	0.416
9-12	*Max*	0.089	0.091	0.039	0.023	2.850	0.860	2.125
	Min	0.006	0.002	0.000	0.000	0.046	0.025	0.975
	Mean	0.031	0.029	0.008	0.007	0.599	0.170	1.311
	SD	0.021	0.019	0.005	0.005	0.521	0.142	0.224

As in the listening and reading domains, tool access per item is fairly limited for the speaking domain (Table 37). Magnifier use per item was an exception across all grades. There are also slight variations among grade-level clusters. For example, activation of the tools per item was most limited among Grade 1 students. On the other hand, item level tool activation was comparatively high among Grades 4-5 and Grades 6-8 students.

On average 2.2 % of the Grade 1 students activated the Magnifier when responding to a speaking item. Activation of the other tools for each item was minimal. For instance, only an average of 0.08% of the Grade 1 students used the Line Guide per item, and only 0.05% clicked on the Highlighter. The students neglected the Color tools in particular, as the mean percentage of students accessing these two tools per item was just 0.007%. Despite overall tool access being limited, particular items triggered a slightly higher percentage of students to activate certain tools.

Similarly, among the Grades 2-3 students, item level tool activation was low. The Magnifier was the most accessed tool at the item level, with an average of 1.95% of students activating it per item. Activation of other tools was just low as it was in the Grade 1. On average, 0.09% of the students used the Highlighter, and 0.08% of the students activated the Line Guide when responding to any given item. Items *14570* and *14572* might be further investigated as a relatively high percentage of students activated a combination of tools while responded to those items, specifically the Highlighter, Line Guide, and Magnifier.

In the Grades 4-5 grade-level cluster, a similar pattern emerged with respect to item level tool activation. Despite being the most accessed tool, the mean percentage of students activating the Magnifier (1.86%) dropped slightly in comparison to activation of that tool by the Grades 2-3 students. The use of other tools was still low overall but relatively high in comparison to use by Grades 2-3 students. On average, 0.16% of the students activated the Highlighter per item and 0.14% used the Line Guide per item. Items *14503* and *14738* could be further examined as a relatively high percentage of students accessed the Highlighter, Line Guide, and Magnifier while responding to these items compared to the other items in the speaking domain.

Among the Grades 6-8 students, Magnifier activation per item was limited to 1.65% on average, which is low compared to access by students in the other grade-level clusters discussed above. Activation of the Magnifier also showed more variation across items in this grade-level cluster compared to the others ($sd = 0.217$). On the other hand, the use of other tools per item was slightly lower and still limited in general. Line Guide and Highlighter activation was limited to about 0.27% of the students on average per item. Item *14505* might be of especial interest for further examination as we observed tool activation, specifically of the Line Guide, Highlighter,

and Magnifier, was at its maximum for this item. Items *14751* and *14752* could also be investigated as tool activation on these items was relatively high for a combination of the tools.

Finally, on average, 1.2% of the Grades 9-12 students activated the Magnifier per item, which was lower than all other grades. Also, on average, 0.23% of the students in the Grades 9-12 grade-level cluster used the Line Guide, and 0.16% used the Highlighter while responding to any given speaking item. Items *14508, 14509,* and *14546* can be further analyzed as a relatively high percentage of students used a combination of the tools on these items compared to other items.

In conclusion, item level tool activation was fairly limited in the speaking domain, with a slight variation among grade-level clusters that is consistent with the domain-level findings. Additionally, although limited, activation of the tools was comparatively high for some items, as can be seen from the maximum values in Table 37.

Table 37. Universal tool use per item in the speaking domain

Grade levels		Color Overlay	Color Contrast	Help (General)	Help (Tools)	Highlighter	Line Guide	Magnifier
1	*Min.*	0.000	0.000	0.014	0.014	0.017	0.035	2.026
	Max.	0.012	0.014	0.062	0.041	0.074	0.148	2.522
	Mean	0.007	0.007	0.039	0.027	0.052	0.088	2.215
	SD	0.003	0.003	0.014	0.009	0.014	0.027	0.163
2-3	*Min.*	0.007	0.005	0.015	0.011	0.059	0.049	1.685
	Max.	0.026	0.025	0.084	0.066	0.160	0.134	2.117
	Mean	0.014	0.013	0.041	0.028	0.094	0.087	1.947
	SD	0.007	0.006	0.019	0.014	0.028	0.028	0.139
4-5	*Min.*	0.008	0.017	0.021	0.011	0.063	0.066	1.678
	Max.	0.088	0.093	0.200	0.130	0.303	0.230	2.079
	Mean	0.045	0.044	0.088	0.066	0.161	0.142	1.864
	SD	0.024	0.024	0.057	0.039	0.072	0.056	0.141
6-8	*Min.*	0.023	0.022	0.019	0.026	0.083	0.088	1.361
	Max.	0.287	0.235	0.488	0.383	0.555	0.640	2.015
	Mean	0.108	0.093	0.182	0.134	0.270	0.277	1.649
	SD	0.070	0.057	0.131	0.095	0.132	0.152	0.217
9-12	*Min.*	0.013	0.004	0.009	0.012	0.044	0.044	0.956
	Max.	0.106	0.086	0.154	0.109	0.317	0.526	1.434
	Mean	0.050	0.039	0.090	0.065	0.160	0.228	1.231
	SD	0.027	0.023	0.051	0.036	0.090	0.135	0.164

Because students in Grades 1-3 handwrite their responses for the writing domain assessment, data on tool use of students in these grades is not available. Similar to other domains, the writing domain analysis (Table 38) showed that activation of the tools per item was low although certain items attracted notably more use.

The mean percentage of Grades 4-5 students accessing the Magnifier was higher than the percentage of students activating any of the other tools, which is consistent with the domain-level findings. Moreover, the average percentage of students using Magnifier (2.6 %), Highlighter (1.5%), and Line Guide (0.8%) was higher than all other grades. Additionally, the percentage of students accessing these tools, in addition to the Sticky Notes, across all items varied more in the Grades 4-5 group than in any other grade-level cluster (i.e., comparison of *sd* across grades). There was also one item (*14729*) for which Grades 4-5 students activated the Highlighter and the Magnifier (4% for both), Sticky Notes (3.6%), and Line Guide (2%). It is highly probable that one or more features of this item caused the high rate of tool activation.

Within the Grades 6-8 group, the average percentage of students activating the Magnifier per item (2 %) was high than the percentage accessing other tools, just as it was in the Grades 4-5 group. The Sticky Notes (1.6 %) and Highlighter (1.3 %) tools were also activated by more students per item, on average, than other tools. Compared to students in other grade-level clusters, a higher percentage of Grades 6-8 students activated Sticky Notes on average per item and the variation was the smallest ($sd = 0.3$). Activation of other tools for each item was fairly limited, especially activation of the Color and Help tools. One particular item triggered more tool access than all other items in the domain. More students accessed universal tools for *Item 14485* than for other items. Specifically, 3.5% of the students used the Line Guide, 3.7% activated the Magnifier, and about 2% used the Highlighter and the Sticky Notes tools.

Finally, although a similar pattern of tool activation per item was observed among Grades 9-12 students as among students in the lower grade-level clusters, the mean percentage of students activating the tools in the Grades 9-12 group was lower than it was in the other grade-level clusters, specifically for the Magnifier (1.5%), Highlighter (0.8%), Line Guide (0.6%), and Sticky Notes (0.6%) tools. As in other grade-level clusters, one specific item (*14478*) presented to the Grades 9-12 students prompted more activation of all of these tools: 1.7% of students used the Line Guide (1.7%), Highlighter (2.4%), Magnifier (2.1%), and Sticky Notes (2.7%).

In conclusion, the average percentage of students activating tools per item was higher in the writing domain than in the listening and reading domains. Additionally, ELs activated four universal tools (the Line Guide, Highlighter, Magnifier, and Sticky Notes) for one item in each grade-level cluster. These three items in the writing domain thus might need further exploration to identify any potential relationship between the items' features and universal tool use.

Table 38. Universal tool use per item in the writing domain

Grade-levels		Color Overlay %	Color Contrast %	Help (General) %	Help (Tools) %	Line Guide %	Highlighter %	Magnifier %	Sticky Notes %
4-5	*Max*	725.000	0.856	0.316	0.188	2.103	4.021	4.100	3.627
	Min	0.092	0.063	0.060	0.023	0.178	0.305	1.651	0.233
	Mean	0.321	0.387	0.193	0.125	0.841	1.495	2.666	1.283
	Sd	0.227	0.282	0.095	0.070	0.662	1.250	0.905	1.179
6-8	*Max*	0.940	1.095	0.227	0.198	1.863	3.724	3.528	1.99
	Min	0.081	0.082	0.037	0.025	0.156	0.207	1.296	1.179
	Mean	0.383	0.452	0.133	0.111	0.730	1.298	2.134	1.569
	Sd	0.301	0.349	0.064	0.054	0.578	1.208	0.822	0.282
9-12	*Max*	0.332	0.294	0.206	0.136	1.671	2.380	2.086	2.742
	Min	0.043	0.047	0.012	0.011	0.110	0.089	1.080	0.070
	Mean	0.135	0.125	0.095	0.066	0.589	0.787	1.494	0.647
	Sd	0.100	0.089	0.071	0.051	0.513	0.823	0.337	0.948

5. Discussion

The purpose of this study was to explore Grades 1-12 ELs' use of accessibility features within the online ACCESS platform. Given the growing prominence of universal design and the increasingly widespread provision of universal tools for fairer and more accessible assessments

for all student populations, the study makes useful contributions to evolving research from several aspects. First, to the authors' knowledge, no study to date has delved into the actual documentation of ELs' universal tool use in K-12 ELP assessments. By addressing this gap in the literature, this study sheds light on the extent to which ELs activate the accessibility features across domains and grade-level clusters, how disability conditions or proficiency levels affect students' activation of the tools, and whether students access these tools more on certain test items. The major findings are recapped below, and follow with a discussion of the implications of these findings.

5.1. The Extent of Tool Use among ELs at Domain Level

Accessibility features are designed to improve access to content and should not inhibit test performance (Crotts-Roohr & Sireci, 2017). The fact that universal tools are present in an assessment does not guarantee their use, and it is important to track students' interaction with these tools to understand real test behavior. Examining the use of accessibility and accommodations features could inform researchers and educators, ultimately increasing the tools' effectiveness in supporting ELs' test-taking experience (Sireci & Faulkner-Bond, 2015) and increasing the relevance of the tools to student test-takers. Abedi (2009) has documented that ELs rarely use the supports presented to them, and they spend more time on the assessment when they take advantage of the available tools. Similarly, Crotts-Roohr and Sireci (2017) observed that only one third of ELs in their study made use of the available tools, and students' reliance on the tools decreased as they progressed through the assessment. Student response time and use of tools were also positively correlated.

The present study yields findings similar to those in the existing literature. A limited number of students activated the universal tools across the four language domains of the

ACCESS assessment. Moreover, the use of most tools was restricted to one or two-time activations. In general, the number of students making use of the tools dropped as students progressed from the listening and reading to the speaking and writing domains, with a few exceptions. In particular, the use of the Color and Help tools notably decreased throughout the test. The decrease in students' used of the Help tools could be attributed to increasing familiarity with the test environment and the tools. The reduction in students' use of the Color tools might indicate that the majority of ELs found the default color and contrast settings of the assessment satisfactory. Sufficient contrast between background and text is important for readability (Liu & Alderson, 2008), and a closer examination of the use of color and contrast tools among subgroups has shown that these tools are serving these groups as intended. For instance, ELs with visual impairments activated of the ACCESS Color and Magnifier tools more than the other tools.

Additionally, ELs were selective in their choice of universal tools. In comparison to Color and Help tools, more ELs took advantage of the Line Guide, Highlighter, and Magnifier across all domains in the assessment. Kim et al. (2018) have found that K-12 language educators also value line guide and highlighter tools more than other features, and they believe that their students frequently use these tools during an assessment. In this study, activation of the ACCESS Highlighter dramatically increased in the reading domain, and half of the students accessed the tool at least nine times. Similarly, in the writing domain, the Highlighter was the most accessed tool, with half of the students activating it at least six times. These findings suggest student selection of the tools could be partially explained by test domain requirements: The reading and writing domains require more text analysis than do other domains, which may have led to the increase in Highlighter use.

Familiarity with the tools might also impact tool selection. Although students are given access to test demos, which describe the universal tools, before they start the assessment, it might not be enough for them to see the actual value of these tools in supporting their test experience. Unless educators emphasize the importance of the tools, ELs may not see the benefit of activating them (Kim et al., 2018). In addition, tools such as Sticky Notes are rare in regular word processors, making them unfamiliar to children. It must also be acknowledged that limited use of a tool does not always signal irrelevance or ineffectiveness. More in-depth information must be gathered from students about their tool choices to be confident about the findings in this study. Specifically, frequent use of the Magnifier and Line Guide tools in the ACCESS speaking domain is not easily explicable and might hint at some disruptive use of the available tools.

The grade level analysis also has noteworthy implications. Like Cohen et al. (2017), use of universal tools varies among different grade-level clusters. While students in Grades 4-5 and Grades 6-8 showed high activation of universal tools, the tools were least appealing to Grade 1 students. According to previous research, when shown on-screen, some tools might interfere with other features or distract students (PARCC, 2017). Cohen et al. (2017) found deteriorating effect on third-graders' performance as a result of including accessibility features. Universal tools in the online ACCESS platform might be disruptive for very young learners. For instance, in the speaking domain, half of the Grade 1 students activated the Magnifier at least 108 times. The novelty of taking an online assessment might have led to meaningless clicks in this section.

5.2. Differential Tool Use between ELs and ELs with Disabilities

Study findings suggest that the universal tools explored in this study could support both ELs and ELs with disabilities, specifically those with visual disabilities and learning difficulties. At least four of the tools—Color Contrast, Color Overlay, Line Guide, and Magnifier—can

potentially support the visual needs of students. The Help tools show how students can navigate the other tools, so they provide support for use of the universal tools. Meanwhile, the Highlighter and Sticky Notes could support the processing and organizational needs of ELs. Results indicate that slightly more ELs with disabilities accessed the universal tools across all domains than did ELs without disabilities. The differences in the percentage of use between the two groups was relatively large in the listening and reading domains, specifically for the Line Guide, Highlighter, and Magnifier. Activation of tools became more similar in the later sections of the test; that is, ELs with and without disabilities were similar in their use of tools across grade-level clusters in the latter (i.e., reading and writing) domains of the test. The two groups showed significant differences with respect to some tool use (e.g., Line Guide, Highlighter, Magnifier). However, differences were attributable to large sample sizes rather than meaningful differences in tool use.

Results also show the varying patterns of tool activation among different disability subgroups. For instance, almost half of the students with visual impairments activated the Magnifier in the listening, reading, and speaking domains. Findings also show that more students with serious emotional disabilities than students with other disability types activated the universal tools across all domains. It is possible that these students just clicked through the tools without actually benefitting from them given their special conditions (i.e., reduced attention span). Additional research is needed to support these types of inferences. However, the present study reveals important findings for ELs with disabilities who are all too often lumped with ELs without disabilities in research.

5.3. The Differential Tool Use among Different Proficiency Levels

The literature has presented mixed findings with respect to the relationship between the use of accessibility features and proficiency levels. Crotts-Roohr and Sireci (2017) found that

intermediate proficiency students used tools more than did students with lower or higher levels of language proficiency. In contrast, in their investigation of the impact of universal tools, Emick and Kopriva (2007) found evidence of the effectiveness of these supports for more proficient students. This study also showed that students at an intermediate and advanced level of proficiency generally displayed more use of the available tools. In this study, there was significant variation in tool use across PLs, especially with respect to the most frequently used tools, such as the Highlighter and Sticky Notes. These findings suggest that universal tools might be more responsive to the needs of students at higher proficiency levels. For students at lower proficiency levels, use of the tools might add to the cognitive load during test taking. Significance testing among the groups showed PL groups differed significantly with respect to use of some tools. Yet all the differences were negligible and were possibly due to large sample sizes.

5.4. The Extent of Tool Use at the Item Level

Although the use of universal tools was more limited at the item level, findings provide some evidence that particular item types triggered more use of tools. In fact, some items stimulated the use of a combination of tools, as discussed above. To understand the nature of the items that triggers increased tool use, future studies should explore the relationships between item characteristics and universal tool use. Cohen et al. (2017) argues that tests of greater difficulty may decrease the effectiveness of supports. Therefore, future research could examine how (a) item difficulty, (b) position of the item in the test, and (c) item type could affect ELs' universal tool use.

5.5. Study Implications

Although no single design can guarantee a test is as accessible as possible for all students (Thompson, Johnstone, Anderson, & Miller, 2005), this study's findings suggest that test developers should be more conscious about the development and inclusion of universal tools in ELP assessments. The accessibility tools available to test-takers should have clear pedagogical and practical values (Kim et al., 2018). Multiple factors should be considered when developing universal tools, such as (a) test-takers' grade level, (b) test-takers' language proficiency, (c) the cognitive load of the test, (d) the number of tools presented, (f) test-takers' familiarity with the tools, and (g) the relevance of the tools to the test content and format. It must be also kept in mind that inclusion of too many universal tools in an online platform might inadvertently lead to the rare inappropriate use of the tools, thereby negatively impacting students' performance (Higgins, Fedorchak, & Katz, 2012).

The present research suggests further optimization of the tools could be beneficial to students. Tools that are geared toward processing information and helping ELs with cognitive strategies (i.e., Highlighter, Sticky Notes) might be more relevant to ELs' needs to understand information presented in the test than manipulate the test environment (i.e., Color tools). In addition, limiting the activation of tools per item and presenting tools depending on the domain requirements might help avoid some of the undesirable outcomes discussed above. There is also a need to bridge assessment and instruction with respect to accessibility tools. Integrating the tools that support test-taking into day-to-day classroom teaching could increase students' familiarity with the tools and the likelihood that students make use of them during testing.

6. Conclusion and Future Directions

Accessibility features present fair test-taking opportunities while supporting students' test-taking experience without compromising the test construct, and thus they have great potential for ELP assessments. The present study aimed to provide evidence for the validity of the universal tools embedded in the online ACCESS test platform by focusing on students' use of the tools. This study explores the validity of the accessibility tools by means of test process data (i.e., telemetry data) and enables direct observation as to whether examinees accessed universal tools, which tools they accessed, and how frequently they accessed those tools. As there is a dearth of research on accessibility for ELs, by focusing on a widely used ELP assessment in the K-12 context, this study has crucial implications for language testing research. The type of process information analyzed, based on telemetry data, not only yields direct information about the activation of accessibility tools, and is also more reliable than solely relying on students' test scores. The present study also makes a deliberate attempt to explore accessibility tools while taking into account different test-taker background variables, such as grade level, disability status, and proficiency level, to provide a more coherent picture of accessibility features use by a heterogeneous EL group.

The present study was exploratory in nature and there are aspects of accessibility feature use that were beyond its scope. In order to address this study's limitations and draw a more complete picture of accessibility features embedded in online ELP assessments, future research should address aspects of universal tool provision such as effectiveness, feasibility of inclusion in a given platform, and differential impact of tool availability and use on various test-taker groups. For instance, the fact that universal tool activation was limited in some domains for some student groups might be attributable to several factors. It would be misleading to conclude that students do not need accessibility tools in responding to certain types of questions. Findings

suggest some students use the tools in line with their needs. In general, a higher percentage of students activated the Highlighter in the reading section than in others. As another example, students with visual impairments made the most use of the Magnifier tool compared to ELs with other types of disabilities. Rather than being viewed as a static property of assessments, accessibility should be considered an interaction between the characteristics of an assessment and a test-taker that enhances or inhibits the ability of the test-taker to respond (Almond et al., 2010). Therefore, it is particularly important to uncover when and under what circumstances students use tools and their perceptions about the available universal tools through cognitive labs or interviews. The degree to which students are familiar with the available tools or the extent to which tools are embedded in classroom instruction might also play an important role in students' ability to effectively apply the tools in the assessment environment. These are just a few of the questions that should be explored during interactions with students.

Another important topic for future research to consider is the cognitive load these universal tools place on students. Although these tools are designed to support the test taking experience, they can introduce additional barriers, especially in terms of the time and effort students must exert to navigate the tools (Crotts-Roohr & Sireci, 2017). Computer-based assessments are more complex and multidimensional in nature than paper-based assessments, and online supports add a further dimension (Abedi & Ewers, 2013). This study suggests some tools might have been distracting for some students, as the range of maximum clicks was quite wide.

Future research should also explore potential issues students could experience in using tools. As the accessibility features and supports intend to increase inclusiveness (Abedi, 2009), studies could investigate variation in test completion rates relative to when these supports are

present or absent. Finally, although the current study tackled the validity of tools with respect to their use and varied use among different subgroups, future studies might consider their impact on actual scores through experimental designs.

References

Abedi, J. (2009). Computer testing as a form of accommodation for English language learners. *Educational Assessment*, *14*(3-4), 195-211.

Abedi, J. (2014). The use of computer technology in designing appropriate test accommodations for English language learners. *Applied Measurement in Education*, *27*(4), 261-272.

Abedi, J. (2017). Utilizing accommodations in assessment. In E. Shohamy et al. (Eds.), *Language testing and assessment, encyclopedia of language and education* (pp. 303-322). New York: Springer.

Abedi., J., Bayley, R., Ewers, N., Mundhenk, K., Leon, S., & Kan, J. (2012). Accessible reading assessments for students with disabilities. *International Journal of Disability, Development and Education*, 59(1), 81-95.

Abedi, J., Courtney, M., & Leon, S. (2003). *Research-supported accommodation for English language learners in NAEP* (CSE Tech. Report No. 586). Los Angeles: The University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Abedi, J., & Ewers, N. (2013). *Accommodations for English learners and students with disabilities: A research based decision algorithm* [Presentations slides]. Smarter Balanced Assessment Consortium. Retrieved from https://portal.smarterbalanced.org/library/en/accommodations-for-english-language-learners-and-students-with-disabilities-a-research-based-decision-algorithm.pdf

Abedi, J. Hofstetter, C., & Lord, C. (2004). Assessment accommodations for English language learners: Implications for policy-based empirical research. *Review of Education Research*, 74(1), 1-28.

Abedi, J., Leon, S., & Kao, J. C. (2008). *Examining differential distractor functioning in reading assessments for students with disabilities* (CRESST Technical Report 743). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Acosta, B. D., Rivera, C., & Shafer Willner, L. (2008). *Best practices in state assessment policies for accommodating English language learners: A Delphi study*. Arlington, VA: The George Washington University Center for Equity and Excellence in Education. Retrieved from https://eric.ed.gov/?id=ED539759

Almond, P., Winter, P., Cameto, R., Russell, M., Sato, E., … Lazarus, S. (2010). Technology-enabled and universally designed assessment: Considering access in measuring the achievement of students with disabilities—A Foundation for Research. *Journal of Technology, Learning, and Assessment*, *10*(5), 1-52.

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

Carr, T. G. (2008, March). *Qualitative review of items that worked and didn't work.* Paper presented at the annual meeting of National Council of Measurement in Education, New York, NY.

Cawthon, S. (2010). Assessment accommodations for English language learners: The case of former-LEPs. *Practical Assessment, Research & Evaluation*, *15*(13), 1-9.

Choi, I., & Cho, Y. (2016, April). *The impact of spellchecker use during an English writing assessment: A case study*. Paper presented at annual meeting of the American Association for Applied Linguistics, Orlando, FL.

Cohen, D., Tracy, R., & Cohen, J. (2017). On the effectiveness of pop-up English language glossary accommodations for EL students in large-scale assessments. *Applied Measurement in Education*, 30(4), 259-272.

Council of Chief State School Officers. (2016). *Major provision of Every Student Succeeds Act (ESSA) related to the education of English learners.* Washington, DC: CCSSO.

Crotts-Roohr, K., & Sireci, S. G. (2017). Evaluating computer-based test accommodations for English learners. *Educational Assessment*, *22*(1), 35-53.

Dolan, R. P., Burling, K. S., Rose, D., Beck, R., Murray, E., … Strain-Seymour, E. (2010). *Universal design for computer-based testing (UD-CBT) guidelines*. Retrieved from http://images.pearsonassessments.com/images/tmrs/tmrs_rg/TMRS_RR_UDCBTGuidelinesrevB.pdf

East, M. (2007). Bilingual dictionaries in test of L2 writing proficiency: Do they make a difference*? Language Testing*, *24*(3), 331-353.

Emick, J., & Kopriva, R. J. (2007, April). *The validity of large-scale assessment scores for ELs under optimal testing conditions: Does validity vary by language proficiency?* Paper presented at the annual meeting of American Educational Research Association, Chicago, IL.

Frankenberg-Garcia, A. (2005). A peek into what today's language learners as researchers actually do. *International Journal of Lexicography*, *18*(3), 335-355.

Frankenberg-Garcia, A. (2011). Beyond L1-L2 equivalents: Where do users of English as a foreign language turn for help? *International Journal of Lexicography*, *24*(1), 97-123.

Guzman-Orth, D., Laitusis, C., Thurlow, M., & Christensen, L. (2016). *Conceptualizing accessibility for English language proficiency assessments* (Research Report No. RR-16-07). Princeton, NJ: Educational Testing Service.

Hansen, E. G., & Mislevy, R. J. (2006). Accessibility of computer-based testing for individuals with disabilities and English language learners within a validity framework. In M. Hricko & S. Howell (Eds.), *Online assessment and measurement: foundations and challenges* (pp. 214-261). Hershey, PA: Information Sciences.

Higgins, J., Fedorchak, G., & Katz, M. (2012). *Assignment of accessibility tools for digitally delivered assessments: Key findings.* White paper for Enhanced Assessment Grant #S368A100008.

Hoffmann, T. (2007). *Final report for universal assessment system phase I research*. Wellesley, MA: Nimble Assessment Systems.

Howell, D. C. (2013). *Statistical methods for psychology*. Belmont, CA: Cengage Learning.

Kim, A., Monroe, M., & Lee, S. (2018). *Examining K-12 educators' perception and instruction of online accessibility tools in language assessments, and English language learners' use of the tools.* Manuscript in preparation.

Kopriva, R. J., Triscari, R., & Carr, T. G. (2014). *Exploring a technology-based, multi-semiotic methodology for measuring challenging content knowledge and skills of low English proficient students*. Manuscript submitted for publication.

Kopriva, R. J., Winter, P. C., Triscari, R., Carr, T. G., Cameron, C., & Gabel, D. (2013). *Assessing the knowledge, skills, and abilities of ELs, selected students with disabilities and controls on challenging high-school science content: results from randomized trials of ONPAR and technology-enhanced traditional end-of-course biology and chemistry tests.* Madison, WI: Institute for Innovative Assessment, University of Wisconsin-Madison. Retrieved from http://iiassessment.wceruw.org/research/researchPapers/Assesing%20KSAs%20ONPAR%20HS%20Exper%20Study%20Results%203%205%2013.pdf

Liu, K. K., & Anderson, M. (2008). Universal design considerations for improving student achievement on English language proficiency tests. *Assessment for Effective Intervention, 33*(3), 167-176.

Messick, S. (1996). *Validity and washback in language testing*. Princeton, NJ: Educational Testing Service.

National Center for Education Statistics [NCES]. (2018). The condition of education: English language learners in public schools [Webpage]. Retrieved from https://nces.ed.gov/programs/coe/indicator_cgf.asp

National Education Association. (n.d.). English language learners. Retrieved from http://www.nea.org/home/32409.htm.

Oh, S. R. (2018). *Investigating test-takers' use of linguistic tools in second language academic writing assessment* (Doctoral dissertation). Retrieved from ProQuest. (10747964).

Partnership for Assessment of Readiness for College and Careers. (2017). *PARCC accessibility features and accommodation manual*. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiSzKaUq63eAhXEY98KHTCwCSkQFjAAegQICBAC&url=http%3A%2F%2Favocet.pearson.com%2FPARCC%2FDocuments%2FGetFile%3FdocumentId%3D4901&usg=AOvVaw3UwXvbIsNDAczs353-Qtew

Pennock-Roman, M., & Rivera, C. (2011). Mean effects of test accommodations for ELs and non-ELs: A meta-analysis of experimental studies. *Educational Measurement: Issues and Practice*, *30*(3), 10-28.

Rivera, C., Collum, E., Shafer Willner, L., & Sia, J. K., Jr. (2006). An analysis of state assessment policies addressing the accommodation of English language learners. In C. Rivera & E. Collum (Eds.), *A national review of state assessment policy and practice for English language learners* (pp. 1–173). Mahwah, NJ: Erlbaum.

Russell, M., Hoffman, T., & Higgins, J. (2009). Meeting the needs of all students: A universal design approach to computer-based testing. *Innovate: Journal of Online Education*, *5*(4), 1-6.

Russell, M., C. Johnstone, J. Higgins, and T. Hoffmann. (2008). *FCAT computer accommodations pilot study final report*. Report prepared for the Florida Department of Education, Tallahassee, FL.

Shafer Willner, L., & Monroe, M. (2016). *The WIDA accessibility and accommodations framework: Considerations influencing the framework development*. Madison, Wisconsin: WIDA.

Shafer Willner, L., Rivera, C., & Acosta, B. (2008). *Descriptive study of state assessment policies for accommodating English language learners*. Arlington, VA: The George Washington University Center for Equity and Excellence in Education.

Shafer Willner, L. & Rivera, C. (2011, Winter). *Are EL needs being defined appropriately for the next generation of computer-based tests? AccELLerate 3*(2), 12-14. . Retrieved from National Clearinghouse for English Language Acquisition website: http://www.ncela.gwu.edu/files/uploads/17/Accellerate_3_2.pdf

Sireci, G. S., & Faulkner-Bond, M. (2015). Promoting validity in the assessment of English language learners. *Review of Research in Education*, *39*, 215-252.

Sireci, S. G., & Wells, C. (2010). Evaluating the comparability of English and Spanish video accommodations for English language learners. In P. C. Winter (Ed.), *Evaluating the comparability of scores from educational achievement test variations* (pp. 33–68). Washington, DC: Council of Chief State School Officers.

Thompson, S. J., Johnstone, C. J., Anderson, M. E., & Miller, N. A. (2005). *Considerations for the development and review of universally designed assessments* (Technical Report 42). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved from https://nceo.umn.edu/publications/OnlinePubs/Technical42.htm

Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large-scale assessments* (NCEO Synthesis Report No. 44). Minneapolis: The University of Minnesota, National Center on Educational Outcomes.

Thurlow, M. L., & Kopriva, R. J. (2015). Advancing accessibility and accommodations in content assessments for students with disabilities and English learners. *Review of Research in Education*, *39*(1), 331-369.

Thurlow, M., Lazarus, S. S., Albus, D., & Hodgson, J. (2010). *Computer-based testing: Practices and considerations* (Synthesis Report 78). Minneapolis, MN: The University of Minnesota, National Center on Educational Outcomes.

U.S. Congress. (2015). *Every Student Succeeds Act*. Washington, DC: Author.

U.S. Department of Education, Office of Special Education and Rehabilitative Services. (2000). *A guide to the individualized education program.* Washington, DC: Author. Retrieved from https://www2.ed.gov/parents/needs/speced/iepguide/iepguide.pdf

U.S. Department of Education. (2014). *Questions and answers regarding the inclusion of English learners with disabilities in English language proficiency assessments and Title III annual measurable achievement objectives.* Washington, DC: Author. Retrieved from https://www2.ed.gov/policy/speced/guid/idea/memosdcltrs/q-and-a-on-elp-swd.pdf

U.S. Department of Education. (2016). *Parent and education guide to section 504 in public elementary and secondary schools.* Washington, DC: Author. Retrieved from https://www2.ed.gov/about/offices/list/ocr/docs/504-resource-guide-201612.pdf

U.S. Department of Education. (2018). *A state's guide to the U.S. Department of Education's assessment peer review process.* Washington DC: Author.

Willner, L. S., & Monroe, M. (2016). *The WIDA accessibility and accommodations framework: Considerations influencing the framework development.* Madison, WI: WIDA Consortium.

World-Class Instructional Design and Assessment [WIDA]. (2018). *2017 - 2018 accessibility and accommodation supplement.* Retrieved from https://wida.wisc.edu/sites/default/files/resource/ACCESS-Accessibility-Accommodations-Supplement.pdf

Wright, L. J., & Kopriva, R. J. (2009). *Using cognitive labs to refine technology-enhanced assessment tasks and ensure their accessibility: Insights from data collected to inform ONPAR elementary and middle school science task development.* Madison, WI: Institute for Innovative Assessment, University of Wisconsin-Madison. Retrieved from http://iiassessment.wceruw.org/research/researchPapers/Insights%20ONPAR%20El%20and%20MS%20Science%20Cog%20Labs%2009b.pdf

Zehler, A. M., Fleischman, H. L., Hopstock, P. J., Pendzick, M. L., & Stephenson, T. G. (2003). *Descriptive study of services to LEP students and LEP students with disabilities*. Report submitted to Office of English Language Acquisition, U.S. Department of Education (Contract No. ED-00-CO-0089). Retrieved from https://www.ncela.us/files/rcd/BE021199/special_ed4.pdf

Suggested Citation

Kim, A., Yumsek, M., Chapman, M. & Cook, H. G. (2019). *Investigating K-12 English Learners' Use of Universal Tools Embedded in Online Language Assessments* (WIDA Technical Report No. TR-2019-2). Madison, WI: WIDA at the Wisconsin Center for Education Research.

Appendix. Distribution of Disability Types in All Domains

	Listening	Reading	Speaking	Writing
Autism Spectrum Disorder	1.29	1.29	1.28	0.99
Deaf-blindness	0.01	0.01	0.01	0.01
Developmental Delay	2.02	2.01	2.03	0.10
Hearing impairment, including deafness	0.31	0.33	0.31	0.33
Intellectual Disability	1.26	1.25	1.22	1.48
Multiple Disability	0.30	0.30	0.30	0.32
Other Health Impairment	3.17	3.17	3.18	3.35
Orthopedic Impairment	0.12	0.12	0.12	0.10
Serious Emotional Disability	0.87	0.85	0.83	1.00
Specific Learning Disability	18.00	17.97	18.06	21.47
Speech or Language Impairment	6.48	6.48	6.56	3.55
Traumatic Brain Injury	0.08	0.08	0.08	0.09
Visual Impairment, including blindness	0.06	0.05	0.05	0.05
NA	66.03	66.09	65.97	67.16