



**A Summary of the  
Research on the  
Effects of K-12 Test  
Accommodations: 2021**

**NCEO Report 438**



**NCEO**  
National Center on  
Educational Outcome

# **A Summary of the Research on the Effects of K–12 Test Accommodations: 2021**

Christopher M. Rogers, Virginia A. Ressa, Sheryl S. Lazarus,  
Martha L. Thurlow, and Ghada Swadek

**February 2023**

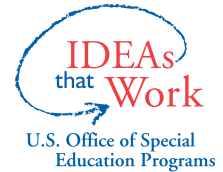
All rights reserved. Any or all portions of this document may be reproduced and distributed without prior permission, provided the source is cited as:

Rogers, C. M., Ressa, V. A., Lazarus, S. S., Thurlow, M. L., & Swadek, G. (2023). *A summary of the research on the effects of K–12 test accommodations: 2021* (NCEO Report 438). National Center on Educational Outcomes.



The Center is supported through a Cooperative Agreement (#H326G210002) with the Research to Practice Division, Office of Special Education Programs, U.S. Department of Education. The Center is affiliated with the Institute on Community Integration at the College of Education and Human Development, University of Minnesota. Consistent with EDGAR §75.62, the contents of this report were developed under the Cooperative Agreement from the U.S. Department of Education, but do not necessarily represent the policy or opinions of the U.S. Department of Education or Offices within it. Readers should not assume endorsement by the federal government.

**Project Officer:** David Egnor



**In collaboration with:**



### **NCEO Core Staff**

Sheryl S. Lazarus, Director  
Jessica Bowman  
Kascinda Fleming  
Gail Ghere  
Linda Goldstone  
Andrew R. Hinkle  
Kristi K. Liu  
Charity Funfe Tatah Mentan

Michael L. Moore  
Darrell Peterson  
Virginia A. Ressa  
Christopher M. Rogers  
Jennifer Sommerness  
Kathy Strunk  
Martha L. Thurlow  
Yi-Chen Wu

National Center on Educational Outcomes  
University of Minnesota  
2025 East River Parkway, Room 1-330 • Minneapolis, MN 55414  
Phone 612/626-1530  
<http://www.nceo.info>

The University of Minnesota shall provide equal access to and opportunity in its programs, facilities, and employment without regard to race, color, creed, religion, national origin, gender, age, marital status, disability, public assistance status, veteran status, sexual orientation, gender identity, or gender expression.

This document is available in alternative formats upon request.

## Executive Summary

---

The relevance and importance of academic research on assessment accommodations continues to be pertinent to students with disabilities and those developing accommodations policies at the state level. The research published in the year 2021 addressed several critical areas such as the effects of accommodations on student performance and educators' knowledge, perceptions, and implementation practices regarding accommodations. The research also identifies emerging trends such as the use of paper-and-pencil accommodations in digital environments, the importance of individualized accommodations in lieu of standard accommodations, and the increased research in the science, technology, engineering, and mathematics (STEM) content areas. These findings and trends provide important and relevant information for policymakers.

The studies presented in this report provide an update on the state of the research on testing accommodations. The National Center on Educational Outcomes (NCEO) has covered research published since 1999. In this report, we summarize the research published in 2021, with 12 research studies addressing testing accommodations in the U.S. K–12 education system.

**Purpose of research:** Several purposes related to accommodations were addressed in the research published in 2021. The most commonly reported primary research purpose was the effects of accommodations on the performance of students (including students with disabilities, students without disabilities, or both). It was a purpose of three-fourths of the studies (n=9). A majority of the studies (83%) had additional purposes which included perceptions, implementation/use, discussion of issues, summarization of the research, comparison of test items, and test validity.

**Research design:** The research published in 2021 included several study designs. Descriptive quantitative (n=3) and quasi-experimental (n=5) research designs were the predominant designs, with these two methods combined comprising close to 70% of the studies. Experimental and descriptive qualitative designs were used in two studies each. Most of the studies (n=11) used primary data collection sources, with only one study using secondary data sources. Tests, surveys, and interviews were used in most studies. Over half of the studies used more than one method to collect data, with survey and test combinations as the most common.

**Types of assessments, content areas:** A variety of assessment types were used to investigate student performance. Surveys or academic tests developed by professionals or researchers using sources outside the study were used in nine studies. Non-academic protocols or surveys developed by study authors made up the next largest group (n=5). Three studies used norm-referenced cognitive ability measures. Course grades, criterion-referenced academic achievement measures, norm-referenced academic achievement measures, and state criterion-referenced assessments were used in one study each. Four studies used other types of instruments.

Nine studies examined specific content areas assessed in this year's report. Eight studies, comprising almost 90 percent of the studies, assessed mathematics. Two studies (22%) each

analyzed reading, science, and other language arts performance. Three studies, comprising 33% of relevant studies, analyzed multiple content areas.

**Participants:** Participants in the studies reviewed included K–12 students. Most of the studies had student participants in the middle school (50%) or high school level (60%). One study had participants who were in kindergarten. In contrast to the research published in 2020, which included more than half of the studies from multiple schooling levels, only one study published in 2021 had participants from more than one schooling level. Educators were participants in two studies as respondents or interviewees. Study sample sizes ranged between three and 481,282 participants.

**Disability categories:** A notable finding in this report on accommodations research published in 2021 is that four studies (33%) did not include students with disabilities and three studies (25%) did not specify the students' disabilities. Participants with disabilities were the subjects of eight studies (67%). This differs from previous years when most study participants were students with disabilities. Students with attention problems were represented in two studies, as were students with learning disabilities. Autism, intellectual disabilities, speech/language disabilities, and visual impairments were each included in one study.

**Accommodations:** The 2021 research studies on accommodations included a variety of accommodation categories. The studies included presentation, equipment and materials, and response accommodations as well as studies with multiple categories of accommodations. No studies addressing scheduling and timing accommodations or setting accommodations were published in 2021. The two most frequently studied accommodations categories were presentation (represented in 67% of the studies) and equipment and materials (represented in 58% of the studies). Electronic administration and the use of digital tools were prevalent across the accommodations categories and studies. In all, the 12 studies reviewed investigated the use of 19 accommodations used by students with and without disabilities.

**Findings:** The research on accommodations published in 2021 included eight studies with findings on the effects of accommodations on student performance. Three studies (Long et al.; Marble-Flint et al.; Zhang et al.) showed positive effects of accommodations, two studies (Bohack; Davis et al.) indicated essentially no difference between the accommodated and non-accommodated assessments, and two studies (Wang et al.; Zebehazy & Wilton) found a negative effect. Finally, one study (Zhang & Rivera) had mixed results.

This review of 2021 accommodations research identified six studies with findings related to student and teacher perceptions of accommodations (Caldone; Davis et al.; Long et al.; Nnoli et al.; Wang et al.; Zebehazy & Wilton). Four studies examined student perceptions of accommodations (Davis et al.; Long et al.; Wang et al.; Zebehazy & Wilton) while two looked at teacher perceptions (Caldone; Nnoli).

## Table of Contents

---

Executive Summary .....	iii
Overview.....	1
Review Process .....	1
Results.....	3
Publication Type .....	3
Research Purposes .....	4
Research Type and Data Collection Source.....	6
Data Collection Methods and Instruments.....	7
Content Area Assessed.....	9
Research Participants .....	10
School Level .....	12
Disability Categories.....	13
Types of Accommodations.....	14
Research Findings.....	15
Effects of Accommodations .....	16
Perceptions about Accommodations .....	17
Implementation and Use of Accommodations.....	19
Discussion .....	21
Report References .....	23
2021 K–12 Accommodation References ( <i>N</i> =12) .....	25
Appendix A: Research Characteristics for K–12 Studies in 2021 .....	27
Appendix B: Research Purposes for K–12 Studies in 2021 .....	28
Appendix C: Instrument Characteristics for K–12 Studies in 2021 .....	32
Appendix D: Participant Characteristics for K–12 Studies in 2021 .....	36
Appendix E: Accommodations Studied for K–12 Studies in 2021.....	38
Appendix F: Findings for K–12 Studies in 2021 .....	42



## Overview

---

Access to assessments for students with disabilities is supported through the use of accommodations. With accommodations, students with disabilities, including English learners with disabilities, are better able to show their academic knowledge and skills. Accommodations also enable these students to participate in state assessments, as required by the Individuals with Disabilities Education Act (IDEA) and by the Elementary and Secondary Education Act (ESEA). Accommodations are changes in materials and procedures that do not compromise the validity of assessment results and interpretations of those results. Evidence is needed to ensure that validity is not compromised. It is also important to examine perceptions of accommodations and implementation issues because these influence whether accommodations are used appropriately. Research conducted on accommodations can provide states with information useful for policy on accommodations.

To synthesize accommodations research efforts completed across the years, the National Center on Educational Outcomes (NCEO) has published a series of reports on accommodations research. The time periods included 1999–2001 (Thompson et al., 2002), 2002–2004 (Johnstone et al., 2006), 2005–2006 (Zenisky & Sireci, 2007), 2007–2008 (Cormier et al., 2010), 2009–2010 (Rogers et al., 2012), 2011–2012 (Rogers et al., 2014), 2013–2014 (Rogers et al., 2016), 2015–2016 (Rogers et al., 2019), 2017 (Rogers et al., 2020), 2018 (Rogers et al., 2021), 2019 (Rogers, Thurlow, et al., 2022), and 2020 (Rogers, Ressa, et al., 2022).

The purpose of this report is to present a synthesis of the research on test accommodations for U.S. elementary and secondary students (K–12) published in 2021. The academic literature described here incorporates empirical studies of performance comparability, as well as investigations into accommodations use, implementation practices, and perceptions of the nature and effectiveness of accommodations. Reporting the findings of recent research studies was the collective goal of these analyses.

## Review Process

---

Similar to the process used in NCEO’s previous accommodations research syntheses, a number of sources were accessed to complete the review of the K–12 accommodations research published in 2021. Specifically, five research databases were consulted: Educational Resources Information Center (ERIC), PsycINFO, Academic Search Premier, Dissertations and Theses Global, and Educational Abstracts. To help affirm the thoroughness of our searches, we used the Web search engine Google Scholar to locate additional research, if any. In addition, a hand-search of at least 50 journals was completed, in efforts to ensure that no qualifying study was missed. A

list of hand-searched journals is available on the NCEO website (<https://nceo.info/Resources/bibliographies/accommodations/methods-for-identifying>).

Online archives of several organizations also were searched for relevant publications. These organizations included the ACT Research Publications (<https://www.act.org/content/act/en/research/reports/act-publications.html>), the College Board Research Library (<http://research.collegeboard.org>), the ETS ReSEARCHER database, and the Wisconsin Center for Educational Research (WCER; <https://www.wcer.wisc.edu/publications>).

The initial search was completed in December 2021. A second search was completed in May 2022, to ensure that all articles published in 2021 were found and included in this review. Within each of these research databases and publications archives, we used a sequence of search terms. Terms searched for this review were:

- standardized (also large-scale, state, standards-based) test (also testing) changes
- standardized (also large-scale, state, standards-based) test (also testing) modification(s)
- standardized (also large-scale, state, standards-based) test (also testing)
- accommodation(s)
- test changes
- test modifications
- test accommodations

Many of these search terms were used as delimiters when searches yielded large pools of documents found to be irrelevant to the searches.

The research documents from these searches were then considered for inclusion in this summary report using several criteria.

1. This analysis included only research published or defended (in doctoral dissertations) in 2021.
2. The scope of the research was limited to investigations of accommodations for regular assessments; hence, studies specific to accommodations for alternate assessments, accommodations for instruction or learning, and universal design in general were not part of this review.
3. Research involving English learners was included only if the target population was English learners with disabilities.



4. Presentations from professional conferences were not searched or included in this review, based on NCEO’s criterion to include only research that would be accessible to readers and which had gone through the level of peer review typically required for publication in professional journals or through a doctoral committee review. (This criterion was implemented for the first time during the 2007–2008 review.)
5. To be included in the online bibliography and summarized in this report, studies needed to involve: (a) experimental manipulation of an accommodation; (b) investigation of the comparability of test scores across accommodated and non-accommodated conditions, or across more than one accommodated condition; or (c) examination of survey results or interview data sets about students’ or teachers’ knowledge or perceptions of accommodations.
6. This report is focused on research on students in United States schools; consequently, studies with only participants in other national contexts were not included.
7. The current report includes only research pertaining to the primary and secondary levels of the education system, that is, from kindergarten through grade 12.
8. We did not include literature reviews or meta-analyses in this review (unlike in NCEO accommodations research reports published prior to 2017).

These criteria do not necessarily apply to NCEO’s Accommodations for Students with Disabilities Bibliography, which is an online database (<https://nceo.info/Resources/bibliographies/accommodations/bibliography>). The Bibliography will continue to include research in non-U.S. settings. In addition, postsecondary accommodations research will continue to be included, and many literature reviews of various kinds have been and will continue to be included in the database as well.

To reflect the wide range of K-12 accommodations research that was published in 2021, the studies were examined and summarized based on the following characteristics: (a) publication type, (b) research purposes, (c) research type and data collection source, (d) data collection methods and instruments, (e) content area assessed, (f) school level, (g) disability categories, and (h) types of accommodations.

## Results

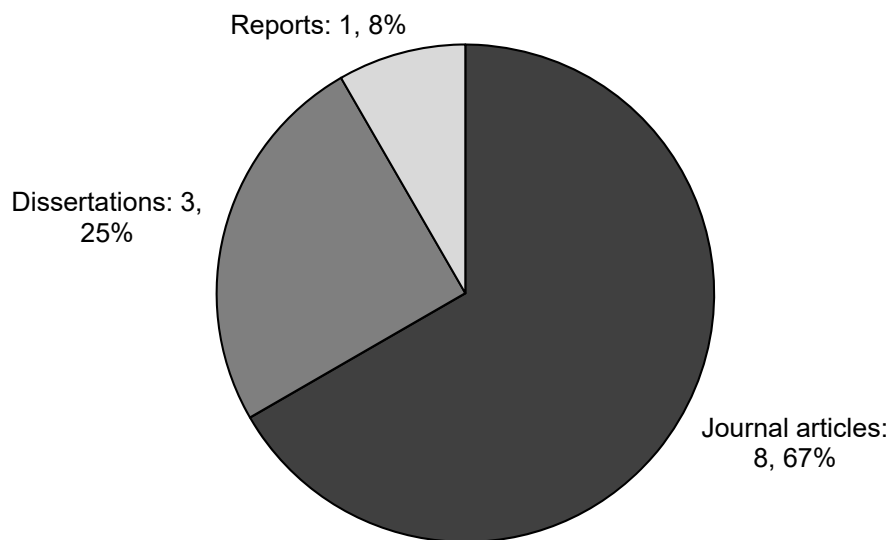
---

### Publication Type

Twelve studies met the inclusion criteria for publications in 2021. Figure 1 shows that the studies included journal articles ( $n=8$ ), dissertations ( $n=3$ ), and a report ( $n=1$ ). Two-thirds of these

studies were published as journal articles, showing a shift from last year, in which there was a nearly even split between journal articles and dissertations. There had been a consistent increase in journal articles in recent years (from four in 2015 to 10 in 2019) but that pattern ended in 2020 when the number of journal articles decreased to six, but once again increased in 2021. The eight journal articles in this year’s review represent eight different academic research journals. One of the 2021 studies (Wang et al.) was a professional report conducted by the research division of ACT. This is the first time since 2015 that an accommodation research study was published by a research organization. Appendix A includes details regarding each study’s publication type.

**Figure 1. Percentage of Accommodations Studies by Publication Type in 2021**



## Research Purposes

The 2021 research on K-12 accommodations had several purposes. The “primary purpose” of each of the 12 studies is shown in Table 1. All of the purposes of the studies are shown in Table 2, including the primary and secondary, or even tertiary, purposes of the studies collectively. We identified the primary purpose of each study based on the narrative description of each study’s research questions, the title of the work, or the first-mentioned purpose in the text.

The primary purpose of nine of the 12 studies reviewed was to compare student performance as a way to analyze the effects of accommodations on student performance. These nine studies included three whose primary purpose was to compare effects for only students with disabilities, four that compared effects for both students with and without disabilities, and two that included only students without disabilities. The balance of studies focused on student and teacher perceptions and preferences about use of accommodations ( $n=2$ ) or implementation practices and accommodations use ( $n=1$ ).

**Table 1. Primary Research Purpose of K–12 Studies in 2021**

<b>Purpose</b>	<b>Number of Studies</b>	<b>Percent of Studies</b>
Compare effects	9	75%
only students with disabilities (3 studies; 25% of studies)		
only students without disabilities (2 studies; 17% of studies)		
both students with and without disabilities (4 studies; 33% of studies)		
Study/compare perceptions and preferences about use	2	17%
Report on implementation practices and accommodations use	1	8%
Discuss issues related to test accommodations	0	0%
Summarize research on test accommodations	0	0%
Compare test items across assessment format	0	0%
Investigate test validity under accommodated conditions	0	0%

The majority of the K-12 research studies on accommodations published in 2021 (10 out of 12 studies) included multiple purposes. Table 2 shows that all research purposes included studying student and teacher perceptions ( $n=6$ ), reporting on implementation practices and accommodations use ( $n=5$ ), discussing issues related to accommodations ( $n=4$ ), summarizing research ( $n=3$ ), comparing test items ( $n=1$ ), and investigating test validity ( $n=1$ ). Because Table 2 includes both the primary and additional purposes from all of the reviewed studies, the total purposes identified exceeds 12.

While just two of the studies focused on student and teacher perceptions as their primary purpose, an additional four studies included student and teacher perceptions as an additional purpose. Similarly, just one study reported on implementation practices and accommodations use as its primary purpose, with four studies including implementation as an additional purpose. Though discussion of issues related to accommodations was not a primary purpose of any of the 12 studies published in 2021, it was a secondary purpose of four studies, two of which were dissertations. Both compared test items and investigated test validity were additional purposes for just one study each. Appendix B includes more details on the purposes of the studies reviewed.

**Table 2. All Research Purposes of K–12 Studies in 2021**

<b>Purpose</b>	<b>Number of Studies</b>	<b>Percent of Studies</b>
Compare effects	9	75%
only students with disabilities (3 studies; 25% of studies)		
only students without disabilities (2 studies; 17% of studies)		
both students with and without disabilities (4 studies; 33% of studies)		
Study/compare perceptions and preferences about use	6	50%
Report on implementation practices and accommodations use	5	42%
Discuss issues related to test accommodations	4	33%
Summarize research on test accommodations	3	25%
Compare test items across assessment format	1	8%
Investigate test validity under accommodated conditions	1	8%

*Note.* Ten of 12 studies had more than one purpose; therefore, numbers total more than the 12 studies represented, and percents total more than 100%.

### Research Type and Data Collection Source

Quasi-experimental and descriptive quantitative research designs were the most frequently used designs in the accommodations research literature published in 2021; together they were used in 67% of the studies ( $n=8$ ). Quasi-experimental designs were used in 42% of the studies ( $n=5$ ) and descriptive quantitative designs were used in 25% of the studies ( $n=3$ ). Experimental and descriptive qualitative designs were used in two studies each. In contrast to studies published in 2020, studies published in 2021 showed a wider representation of study types. No studies in either year used a longitudinal design. (See Table 3; Appendix A presents research types and data collection sources for individual studies.)

**Table 3. Research Type and Data Collection Source for K–12 Studies in 2021**

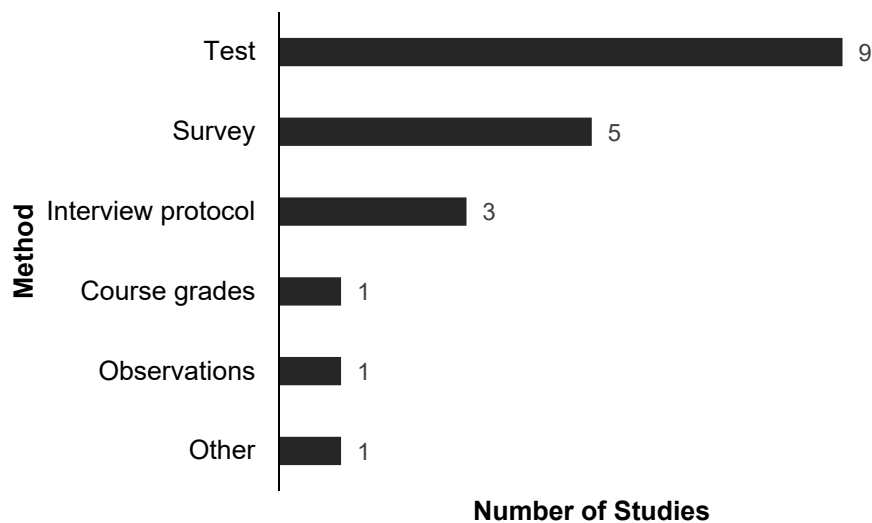
<b>Research Type</b>	<b>Primary Source</b>	<b>Secondary Source</b>	<b>Total</b>
Quasi-experimental	5	0	<b>5</b>
Descriptive quantitative	2	1	<b>3</b>
Experimental	2	0	<b>2</b>
Descriptive qualitative	2	0	<b>2</b>
<b>Total</b>	<b>11</b>	<b>1</b>	<b>12</b>

Most of the studies published in 2021 used primary data collection sources. Only researchers in one study used a secondary data source. Just 8% of the studies used secondary data sources, as compared to 2020 when 27% of studies used secondary data (Rogers, Ressa, et al., 2022), and 2019 when 36% of studies used secondary data (Rogers, Thurlow, et al., 2022).

### Data Collection Methods and Instruments

The 2021 research we analyzed employed the methods shown in Figure 2 to collect study data. Most of the studies ( $n=9$ ) used performance data acquired through academic testing (Bohack; Davis et al; Gilbert et al.; Long et al.; Marble-Flint et al.; Wang et al.; Zebehazy & Wilton; Zhang et al.; Zhang & Rivera). Surveys supplied data for five of the studies (Bohack; Caldone; Davis et al.; Wang et al.; Zebehazy & Wilton). Four studies surveyed students (Bohack; Davis et al.; Wang et al.; Zebehazy & Wilton), and one study (Caldone) surveyed teachers. Interviews were employed in 25% of the studies; interviews were used to gather individual participants’ demographic characteristics as part of broader data-gathering processes with teachers (Caldone; Nnoli) or with students (Long, et al.). Nearly 60% of the 12 studies reported using more than one method or tool to gather data. The most frequently used combination of methods was the combined use of tests and surveys ( $n=4$ ). See Appendix A for additional details about each study’s data collection methods.

**Figure 2. Data Collection Methods Used in All K–12 Studies Published in 2021**



*Note.* Of the 12 studies reviewed for this report, seven reported using two or more data collection methods. Thus, the number of methods in this figure totals more than 12.

Multiple types of data collection tools were employed across the 12 studies on accommodations reviewed in 2021. As shown in Table 4, the most commonly used instruments included surveys or academic tests developed by researchers or other education professionals using sources *outside* of the study. Examples of surveys employed in 2021 studies included the Teachers’ Conception of Assessment (TCoA; Brown, 2006), used by Caldone, and selected items from the student survey administered nationwide to ACT participants that pertained to digital assessment tools (Wang et al.). An example of an academic test developed by professionals or researchers outside of the study is Marble-Flint and colleagues. They used the Preschool Language Assessment Inventory-second edition (PLAI-2; Blank et al., 2003) as a screener for kindergarten study participants. Another example of a tool developed outside of a study is Zhang and Rivera’s application of a set of five test-like tasks—some of which were drawn from Maher and Yankelewitz (2017)—for assessing math fraction sense.

**Table 4. Data Collection Instrument Types for All K–12 Studies Published in 2021**

Instrument Type	Number of Studies <sup>b</sup>	Percent of Studies <sup>b</sup>
Surveys or academic tests developed by professionals or researchers using sources outside of current study	9	75%
Non-academic protocols or surveys developed by study author/s	6	50%
Norm-referenced cognitive ability measures	3	25%
Course grades	1	8%
Criterion-referenced academic achievement measures	1	8%
Norm-referenced academic achievement measures	1	8%
State criterion-referenced assessments	1	8%
Other <sup>a</sup>	4	32%

<sup>a</sup> Other: see Appendix C, Table C-1 for specific information in Lee et al., 2021; Long et al., 2021; Wang et al., 2021; Zebehazy & Wilton, 2021.

<sup>b</sup> Seven studies (58%) used more than one type of instrument; therefore, numbers total more than the 12 studies represented, and percents total more than 100%.

Six studies published in 2021 used non-academic protocols developed by the study authors. These tools included surveys, interview questions or prompts, and observation procedures—that is, protocols (i.e., pre-planned steps) that researchers followed during data collection. Study authors gathered survey responses in three ways: (a) from student participants reporting on their prior familiarity with answering questions using computer software and tools (Bohack); (b) from students expressing their perceptions and preferences about versions of response accommodations (Davis et al.); and (c) from students rating the degree of difficulty of tasks when using different graphic formats (Zebehazy & Wilton). Surveys provided social validity data in three studies (Davis et al.; Wang et al.; Zebehazy & Wilton). Interview protocols designed by study authors were applied in two studies (Long et al.; Nnoli). Long and colleagues gathered first-person data related to social validity by asking students to describe their experiences with

various versions of manipulatives that were provided during the study. Nnoli elicited special education teachers' perceptions of prior experiences when preparing students to participate in tests. Teachers shared their perceptions of actions they took to support students, assessment challenges and how they managed them, and their reflections on themselves currently as well as early in their careers. In the one study that involved observations (Long et al.), researchers documented students' task completion time, rated the degree of task support needed, recorded the level of independent problem solving, and identified how many prompts were needed, if any, and the level of prompts used.

Three norm-referenced cognitive ability measures were administered in three of the studies (Gilbert et al.; Long et al.; Marble-Flint et al.). Two studies (Gilbert et al.; Long et al.) used the Wechsler Intelligence Scale for Children, fifth edition (WISC-V; Wechsler, 2014). Long and colleagues also gathered data using the Wechsler Preschool and Primary Scale of Intelligence, third edition (WPPSI-III; Wechsler, 2002). In one study, Marble-Flint and colleagues gathered aptitude scores on language competence and potential impairments of student participants from the Peabody Picture Vocabulary Test, fourth edition (PPVT-4, Dunn & Dunn, 2007). Zhang and colleagues were the only researchers who made use of student participants' course grades. They used course grades to screen potential participants for the study, selecting those students who had an average grade of 'C' or below in a high school geometry course. A national dataset from a criterion-referenced academic achievement measure administered in the 2019–2020 school year—the ACT in English, mathematics, reading, and science—was analyzed as a primary dependent variable in one study (Wang et al.). One norm-referenced academic achievement measure, the KeyMath3™ diagnostic assessment (KeyMath-3; Connolly, 2007), was employed in one study (Long et al.) for identifying student participants' math grade-equivalence scores. One study (Zhang et al.) used a state criterion-referenced assessment for screening student participants. Zhang and colleagues identified students with poor math performance as study participants through scores on a state math assessment.

Data collection instruments developed by the researchers conducting the study were used in four studies (Lee et al.; Long et al.; Wang et al.; Zebehazy & Wilton). Approximately 58% of all studies ( $n=7$ ) used more than one kind of instrumentation. A complete listing of the instruments used in each of the studies is presented in Table C-1 in Appendix C. The table includes the related studies or other sources for these instruments, when available.

### Content Area Assessed

Ten of the studies published in 2021 focused on accommodations used in specific academic content areas, including cognitive skills. Neither Caldane nor Nnoli identified a specific content area.



As shown in Table 5, mathematics was the most commonly studied content area for studies from 2021. Table 5 also provides the findings for the two previous years (2019—Rogers et al., 2022; 2020—Rogers, Ressa et al., 2022). In all three years, mathematics was the most common content area for accommodations research, though mathematics increased from 50% in 2019 to 80% in 2021. See Appendix C, Table C-2 for details about content areas studied in 2021.

**Table 5. Academic Content Area Assessed in Relevant K–12 Studies across Three Reports**

<b>Content Area Assessed<sup>a</sup></b>	<b>2019</b>	<b>2020</b>	<b>2021</b>
Mathematics	5 (50%)	7 (70%)	8 (80%)
Reading	4 (40%)	5 (50%)	2 (20%)
Science	1 (10%)	1 (10%)	2 (20%)
Writing	1 (10%)	0 (0%)	0 (0%)
Social studies	0 (0%)	0 (0%)	0 (0%)
Other language arts	2 (20%)	0 (0%)	2 (20%)
Cognitive skills	1 (10%)	1 (10%)	1 (10%)
Multiple content	4 (40%)	3 (30%)	3 (30%)
<b>Total (of relevant studies)<sup>b</sup></b>	<b>10</b>	<b>10</b>	<b>10</b>

<sup>a</sup> Studies in all three years included studies that addressed more than one content area (i.e., two, three, or four content areas), so the percentages for each year total more than 100%.

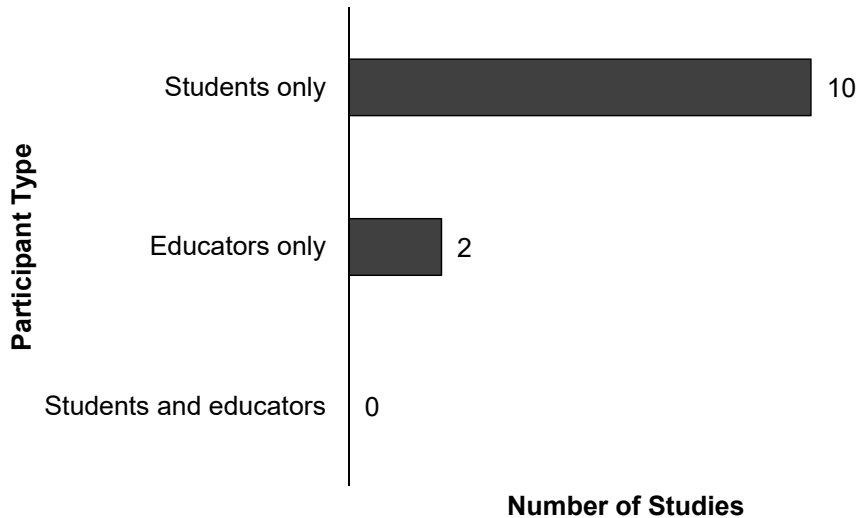
<sup>b</sup> These totals were less than the total number of studies analyzed from these years since not all studies addressed a specific content area; in 2021, two studies (Caldone; Nnoli) did not address a specific content area.

In the research published in 2021, just three of the studies included more than one content area. Wang and colleagues included four content areas: mathematics, reading, other language arts, and science. Lee and colleagues examined data from both math and other language arts, and Zebehazy and Wilton included STEAM content: science, technology, engineering, arts, and math; we documented the content as math and science. None of the studies published in 2021 addressed writing or social studies.

### Research Participants

The studies in this review of accommodations research from 2021 included either students or educators as participants, but none had both students and educators as participants (see Figure 3 and Appendix D). Most of the studies included students only ( $n=10$ , 83%) while a couple studies included only educators ( $n=2$ , 17%).

**Figure 3. Types of Research Participants for All K–12 Studies in 2021**



In 2021, seven studies (Bohack; Lee et al.; Long et al.; Marble-Flint et al.; Zebehazy & Wilton; Zhang et al.; Zhang & Rivera) included students with disabilities in the participant samples (see Appendix D). Two of these seven studies did not specify students' disability category (Bohack; Lee et al.). One study did not provide any information about student participants' disabilities (Wang et al.). Two studies (Davis et al.; Gilbert et al.) included only students without disabilities. Of the remaining studies, two included only educators as participants (Caldone; Nnoli).

The size and composition of the participant groups are shown in Table 6 for the 10 studies in 2021 that included K–12 students. Student participant groups varied from a small sample size of just three participants in one study (Long et al.) to 481,282 participants (Lee et al.). The most common student sample sizes were between 25 and 49 participants ( $n=3$ ) and between 50 and 99 participants ( $n=2$ ). Two studies used extant data sets: one study (Wang et al.) in the 1,000–99,999 participant range, and the other (Lee et al.) in the 100,000–500,000 participant range.

Students with disabilities comprised 0–24% of study samples in four studies (Bohack; Davis et al.; Gilbert et al.; Lee et al.). They comprised 50–74% of samples in two studies (Marble-Flint et al.; Zhang et al.), and 75–100% of the sample in three studies (Long et al.; Zebehazy & Wilton; Zhang & Rivera). Compared with previous reports, more accommodations research studies in 2021 had only students without disabilities or had a smaller proportion of students with disabilities in the sample. Four of nine studies in 2021 included samples with 0% to 24% students with disabilities and five out of nine had samples that had greater than 25% students with disabilities. In comparison, the 2020 report included two out of 10 studies with participant

samples of 0% to 24% students with disabilities and eight out of 10 studies had samples composed of more than 25% students with disabilities (Rogers, Ressa et al., 2022).

**Table 6. Student Participant Sample Sizes and Ratio of K–12 Students with Disabilities in 2021**

Number of Student Participants by Study	Number of Studies by Proportion of Sample Comprising Students with Disabilities				
	0–24%	25–49%	50–74%	75–100%	Total Studies
1–9	0	0	0	1	1
10–24	0	0	0	1	1
25–49	0	0	2	1	3
50–99	2	0	0	0	2
100–999	1	0	0	0	1
1,000–99,999					1 <sup>a</sup>
100,000–500,000	1	0	0	0	1
<b>Total</b>	<b>4</b>	<b>0</b>	<b>2</b>	<b>3</b>	

*Note.* The total number of studies in this table is less than the total number of studies included in this report as two studies (Caldone; Nnoli) only included educators.

<sup>a</sup> One study (Wang et al.) included 16,580 participants, but the percent of students with disabilities was not reported.

Two studies (Lee et al.; Marble-Flint et al.) compared the performance of students with and without disabilities. Two studies (Davis et al.; Gilbert et al.) did not include any students identified with disabilities. Two studies (Bohack; Lee et al.) reported that participants with disabilities were included, but the categories of disabilities were not specified. Wang and colleagues did not specify whether students with disabilities were part of the sample nor did they report any disability categories; therefore, the study was not tallied in the percentages but counted in the total number of studies in Table 6. See Appendix D for additional details about each study’s participants.

### School Level

As shown in Table 7, this review of accommodations research from 2021 identified 10 studies that included participants who were students in the elementary, middle, or high school levels. Most studies with student participants ( $n=8$ ) drew these students solely from within one school level. Some studies had students from a single grade, such as grade 8 (Long et al.), or from multiple grades within a school level, such as grades 6, 7, and 8 (i.e., middle school; Zhang & Rivera). Two studies (Gilbert et al.; Zebehazy & Wilton) included student participants from multiple school levels (i.e., elementary, middle, and high school levels). See Appendix D for students’ specific grade levels when available.

Most of the studies from 2021 investigated accommodations provided to students in high school or middle school. Six studies (Bohack; Davis et al.; Gilbert et al.; Wang et al.; Zebehazy & Wilton; Zhang et al.) had high school participants. Five studies had middle school participants (Gilbert et al.; Lee et al.; Long et al.; Zebehazy & Wilton; Zhang & Rivera). Two studies (Gilbert et al.; Zebehazy & Wilton) had participants from elementary school, and only one study (Marble-Flint et al.) had kindergarten-level participants (i.e., children aged 4–6).

**Table 7. School Level of Student<sup>a</sup> Participants for Relevant K–12 Studies in 2021**

School Level	Number of Studies <sup>b</sup>	Percent of Studies <sup>b</sup>
Kindergarten (K)	1	10%
Elementary school (K–5)	2	20%
Middle school (6–8)	5	50%
High school (9–12)	6	60%

<sup>a</sup> This table includes data for 10 studies. Two studies had only educators as participants and were not incorporated in this table.

<sup>b</sup> Two studies had participants in more than one schooling level; therefore, the counts total more than the 10 studies represented and percents total more than 100%.

## Disability Categories

The accommodations research published in 2021 was composed of students with disabilities from an array of disability categories (see Appendix D for individual study details). As Table 8 shows, the studies included students from six disability categories, with three studies not specifying the disability category. Students with attention problems were included in two studies (Zhang et al., Zhang & Rivera). Similarly, students with learning disabilities were also represented in two studies (Zhang et al., Zhang & Rivera). Most disability categories were represented by participants in just one study. Students with autism (Marble-Flint et al.), intellectual disabilities (Long et al.), speech/language impairments (Zhang & Rivera), and visual impairments (Zebehazy & Wilton) each appeared in one study. Three studies (Bohack; Lee et al.; Wang et al.) did not specify the disability categories for the participating students. Four studies included students without disabilities (33%), with one study comparing performance between students with and without disabilities (Marble-Flint et al.), one study comparing item level accommodation use for students with and without disabilities (Lee et al.), and two studies having no students with disabilities (Davis et al.; Gilbert et al.). For two studies, disability categories did not apply (Caldone; Nnoli) because educators were participants.

**Table 8. Disability Categories of Participants for All K–12 Studies in 2021**

Disabilities of Research Participants	Number of Studies	Percent of Studies
Attention problems	2	17%
Learning disabilities	2	17%
Autism	1	8%
Intellectual disabilities	1	8%
Speech/Language impairments	1	8%
Visual impairments (including blindness)	1	8%
No disability	4	33%
Not specified	3	25%
Not applicable <sup>a</sup>	2	17%

*Note.* The total number of studies is more than the number included in the report because some studies had multiple disability categories.

<sup>a</sup> Educators were the study participants.

## Types of Accommodations

The accommodations categories examined in the 12 research studies published in 2021 are summarized in Table 9. Some studies included accommodations from more than one category (e.g., presentation, response). Presentation accommodations were the most frequently studied category of accommodations ( $n=8$ ). Within the presentation category, three studies addressed electronic administration (Bohack; Gilbert et al.; Marble-Flint et al.). Text-to-speech devices or software were addressed in two studies (Lee et al.; Marble-Flint et al.). Seven additional presentation accommodations were studied by just one study each. See Appendix E, Table E-2 for studies of presentation accommodations.

**Table 9. Accommodation Categories for All K–12 Studies in 2021**

Accommodations Category	Number of Studies <sup>a</sup>
Presentation	8
Equipment and materials	7
Response	4
Scheduling and timing	0
Setting	0

<sup>a</sup> Six studies investigated accommodations from more than one category; therefore, the numbers in this table total more than the 12 studies published in 2021.

The second most common category of accommodations was equipment and materials ( $n=7$ ). Specifically, electronic administration was considered in three studies (Bohack; Gilbert et al.; Marble-Flint et al.). An additional eight equipment and materials accommodations were studied by one study each. As shown in Appendix E, Bohack experimented with three materials accommodations as part of a “bundle” or aggregated set of accommodations used by students.

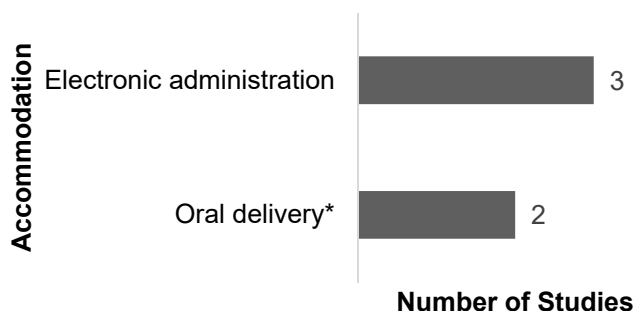
See Appendix E, Table E-3 for additional information about equipment and materials accommodations studies.

Finally, response accommodations were incorporated into four studies. Similar to the previous two categories, electronic administration was a factor in three studies (Bohack; Gilbert et al.; Marble-Flint et al.). The fourth study (Davis et al.) that included accommodations from the response category looked at the use of a digital stylus and keyboard as response devices compared to paper and pencil. See Appendix E, Table E-4 for studies of response accommodations.

As shown in Table 9, none of the accommodations research studied in 2021 included accommodations from the scheduling and timing, or setting categories.

Figure 4 shows the accommodations that were addressed by more than one study published in 2021. Across the studies, we identified 20 accommodations that were investigated, but only electronic administration and oral delivery were included in more than one study. This is different from the studies published in 2020 in which eight accommodations were included in more than one study. The most frequently studied accommodation in 2021 was electronic administration, which was included in three studies (Bohack; Gilbert et al.; Marble-Flint et al.). Oral delivery, which includes live and pre-recorded human voice, was addressed by two studies (Lee et al.; Marble-Flint et al.). Details of the 20 accommodations examined in each of the 12 studies published in 2021 are provided in Appendix E, Table E-1.

**Figure 4. Specific Accommodations Included in More than One K–12 Study in 2021**



*Note.* Accommodations included in only one study are not included in this figure.

\* Oral delivery includes live/in-person and recorded human voice.

## Research Findings

---

The findings of the studies on accommodations published in 2021 are summarized here according to their attributes. These findings were consistent with the stated purposes and focuses of the studies. The findings included sets of research about specific accommodations such as electronic administration. We present findings on the effects of specific accommodations or of

aggregated sets of accommodations during assessments. We also present findings on the impact of other accommodations examined in only one study—such as responding with a digital stylus or keyboard. This section includes findings on the perceptions of accommodations, including those from student test takers and from educators. This section also includes descriptions of implementation conditions as well as patterns of use for various accommodations. In Appendix F, we report findings from individual studies.

## Effects of Accommodations

Of the research published in 2021, eight studies investigated the effects of accommodations on student assessment performance (see Appendix F for details about each study). Three studies (Long et al.; Marble-Flint et al.; Zhang et al.) showed positive effects of the accommodations studied, two studies (Bohack; Davis et al.) found no significant difference between the accommodated and non-accommodated assessments, and two studies (Wang et al.; Zebehazy & Wilton) found negative performance effects. Finally, one study (Zhang & Rivera) yielded mixed results. Although the effects of the two accommodations (i.e., bolded words; written explanations) appeared relatively small, Zhang and Rivera noted that analysis of the students' written explanations of their thinking processes for arriving at their answers suggested that some students overcame barriers on some items due to their use of each of the accommodations.

Three studies from 2021 (Bohack; Gilbert et al.; Marble-Flint et al.) examined the effects of accommodations available through electronic administration. Bohack investigated the effects of an aggregated set of supportive tools embedded into the electronic administration of math assessments for high school students with unspecified disabilities. The potential boost in math performance was analyzed through comparisons to student performance when they completed similar math assessments in a traditional paper-and-pencil format. The embedded electronic tools were designed to support test takers in organizing their thinking when completing test items through features allowing students to eliminate answer options, flag questions for additional attention, take notes using a virtual notepad, and apply a review tool for self-identified items. Bohack found that math performance was not significantly different when completing tests in a digital format, in comparison to a paper-and-pencil format, even with the embedded support tools. Bohack concluded that the math tests were equivalent in that they tested the same academic content across test administration formats. In a second study, Marble-Flint and colleagues analyzed reading comprehension data from kindergarten students with and without autism. The study compared student performance under three different presentation formats or testing conditions. The three testing conditions were adult reading aloud from a storybook in paper format, adult reading aloud from a storybook in an electronic tablet format, and digitally-embedded human voice accompanying the storybook in electronic tablet format; in all three conditions, students replied orally to the questions and their answers were video-recorded and documented with researcher notes. All student participants, with and without autism, performed significantly



better when presented the electronic version of the storybook on a digital tablet, in comparison to the display of the storybook in paper format. Further, students did not score differently when the oral delivery was presented live in-person or through a digitally-embedded human voice recording. The kindergarten participants with disabilities did not differentially benefit from the electronic presentation of the text segments in comparison to their peers without disabilities.

The third study published in 2021 that examined electronic administration was conducted by Gilbert and colleagues. This quasi-experimental study investigated the effects of digital versus paper-based test format on composite score and subscale score on the Wechsler Intelligence Scale for Children, fifth edition (WISC-V; Wechsler, 2014) for 65 students without disabilities. This assessment was administered using an electronic tablet with a touch screen. Student scores on the digital format were generally higher than the standard paper format, indicating that the two formats of the assessment were not equivalent. Researchers suggested that the higher student scores on the electronic format could possibly be attributed to increased student engagement and motivation.

Seven studies examined the effects of accommodations that were each examined in only one study published in 2021. The seven studies and the accommodations examined were:

- Davis et al.: digital stylus, keyboard
- Long et al.: concrete manipulatives, virtual manipulatives
- Marble-Flint et al.: live adult human reader, digitally-embedded recorded human voice
- Wang et al.: paper/pencil, digital
- Zebehazy & Wilton: tactile graphics, tactile graphics with enlarged font
- Zhang & Rivera: annotations, practice items
- Zhang et al.: visual chunking with reference guide, visual chunking without reference guide

Findings for these accommodations are reported in Appendix F.

### Perceptions about Accommodations

This review of accommodations research published in 2021 identified six studies with findings related to student and teacher perceptions of accommodations (Caldone; Davis et al.; Long et al.; Nnoli et al.; Wang et al.; Zebehazy & Wilton). Four studies examined student perceptions of accommodations (Davis et al.; Long et al.; Wang et al.; Zebehazy & Wilton) while two looked at teacher perceptions (Caldone; Nnoli). None of the 2021 studies reported on perceptions of both

students and teachers. This is different from 2020, when there were six studies that reported on perceptions, and two included perceptions of both students and teachers (Rogers, Ressa et al., 2022).

The two 2021 studies in this review that incorporated teachers' perceptions were dissertations by Caldane and Nnoli. Caldane's study of how teachers' beliefs were linked to their assessment and accommodations practices examined middle and high school teachers' perceptions of providing accommodations during classroom and state assessments. Teachers were confident about their ability to provide accommodations across differing classroom assessment tasks and their ability to incorporate accommodations into classroom end-of-course assessments. However, they indicated that it often was difficult to ensure students' access to accommodations for the state standardized assessments. Teachers also expressed difficulty communicating with parents about the purpose of accommodations, and balancing student needs and accommodations in relation to policies.

Nnoli investigated special education teachers' experiences preparing students for high-stakes assessments and providing students with accommodations. Participating teachers reported that it is important to have knowledge of accommodations in students' individualized education programs (IEPs). Teachers also believed that it was important to teach students about the available accommodations and how to effectively use them.

Of the four studies reviewed that considered student perceptions, three were either quasi-experimental or experimental (Davis et al.; Long et al.; Zebehazy & Wilton). These three studies all compared effects of accommodations on student performance (see Appendix B for more details on study purposes). Davis and colleagues compared high school student performance on mathematics assessments using three response options: paper and pencil, digital stylus, and keyboard. This study included a survey to collect information on student perceptions of the three response methods. Students expressed a preference for the paper-and-pencil option, with most students reporting that it was easier to make changes when they used a pencil and could erase. They also thought that their handwriting was clearer with the pencil. The digital stylus was the second most preferred option, and the keyboard was least preferred. Long and colleagues also studied the effects of accommodations on student performance in mathematics. This experimental study compared the effects of digital and concrete algebra manipulatives on student learning, as well as student perceptions and preferences concerning the two types of manipulatives. Manipulatives in both formats (concrete and digital) were equally effective in supporting student learning of algebra content. The small sample of three grade 8 students with intellectual disabilities indicated that they liked using the manipulatives, with a slight preference for the concrete tiles. The choice of concrete or virtual manipulatives did not have a strong impact on student performance, which suggested that the selection of the format of algebra titles as an accommodation should be based on student perceptions and preferences.

Similarly to Davis and colleagues, Zebehazy and Wilton conducted a quasi-experimental study of the effects of accommodations on student performance and collected data about student perceptions. Zebehazy and Wilton compared the performance of students with visual impairments on graphics-based tasks in two presentation accommodation formats (tactile graphics and print graphics, with both including enlarged print). Student performance using either graphic format generally aligned with their teachers' ratings of student ability (e.g., students rated as highly independent in problem solving performed well using both formats). Users of the tactile graphics took longer to complete the tasks (45 minutes for tactile users compared to 28 minutes for print users), suggesting to the researchers that students would have benefited from more practice with the tactile graphics during classroom instruction. Students did not perceive a significant difference in the difficulty of each graphic format. However, researchers found that the difficulty level perceived by students did not match their actual performance. For example, students perceived the bar graph task as low to moderately difficult but none of the tactile graphics users answered all of the items correctly.

Wang and colleagues also studied the effects of testing formats on the performance of high school students. This descriptive quantitative study included analysis of both assessment and survey data to report on students' experiences with a nationally recognized achievement test (ACT) in paper-and-pencil and online formats. Wang and colleagues found that the large majority (94%) of students who completed the test in the online format found an on-screen timer to be helpful for pacing. Students felt that they had enough time to complete the test in both formats. While students perceived that they were not rushed on the English portion of the assessment, researchers found that the English test was the one on which students were flagged for speededness (rushing) most often.

## Implementation and Use of Accommodations

The 2021 studies included five pertaining to accommodations implementation and use (Caldone; Lee et al.; Nnoli; Zhang et al.; Zhang & Rivera). Two studies focused on implementation issues (Caldone; Nnoli), two studies focused on use (Zhang et al.; Zhang & Rivera), and one study focused on both implementation and use (Lee et al.).

Caldone (2021) reported on middle and high school teachers' perceptions and knowledge on accommodation implementation practices. Findings showed that teachers thought they could competently provide accommodations on classroom assessments. When specifically asked about accommodation practices, survey results indicated relatively strong agreement (4=agree; 6=strongly agree) on the application of accommodation practices such as consistently providing accommodations across all assessments (mean=4.97) and providing accommodations on end-of-course assessments (mean=4.96). Teachers' responses on other related practices were not as strong, such as communicating to parents the purpose of accommodations (mean=3.86) or the

relation of student needs and accommodations to policies (mean=3.15). Interview responses related to practices such as differentiating assessments were associated with perceptions of the importance of accommodations. On state assessments, teachers indicated difficulty in ensuring assessment access for students through accommodations. Nnoli (2021) presented special education teachers' perspectives on the implementation of accommodations. Teachers indicated that knowledge of students' IEP accommodations is very important for teachers and students, and that students should be explicitly taught what their accommodations are and how to use them. Further, teachers reported that accommodations were important for students to demonstrate their knowledge and therefore it was important to ensure students receive classroom and testing accommodations. Participating teachers reported spending time preparing students for using accommodations like extended time, speech-to-text software, human reader, calculator tool, voice or immersion reader, and scribe. Findings across the two studies (Caldone; Nnoli) showed that teachers thought they were competent in teaching and implementing accommodations, believed that accommodations are important for students with disabilities to demonstrate their knowledge, and had concerns about students receiving appropriate accommodations to access state assessments.

Two studies (Zhang et al.; Zhang & Rivera) focused on students' use of accommodations in mathematics. Examined together, these studies provided insights into accommodation use. Zhang and colleagues investigated accommodations use by students with math difficulties—including students with low classroom math performance, students who failed state math assessments, and students with disabilities—and by students who were merely at risk of failure in math. As mentioned earlier in this report and detailed in Appendix F, Zhang and colleagues found that both students at risk of failure and students with documented math difficulties benefited from visual chunking—that is, organization of content in sets of information. Only students with math difficulties, including students with learning disabilities, showed significant benefits in geometry performance from both visual chunking and reference guide accommodations. The researchers asserted that the accommodations reduced students' cognitive load by providing cognitive scaffolding, a process that was shown to be specifically beneficial for students with limited math knowledge and skill. The effects of accommodations were moderated by the factors of problem complexity and students' prior geometry knowledge. The second study (Zhang & Rivera) extended Zhang and colleagues' analysis of the impact of accommodations on math performance. Zhang and Rivera's participants completed three versions of a test on fractions: one without supports, one with annotations as an accommodation, and one with whole number warm-up problems as an accommodation. Finding limited performance differences among the test conditions, the researchers concluded that the effectiveness of accommodations might not best be demonstrated with performance score differences, due to individual students' unique difficulties in overcoming obstacles in problem solving. Further, Zhang and Rivera argued the general use of a standard protocol of accommodations for all students ought to be set aside in favor of targeting individualized sets of accommodations for students with math disabilities.

They demonstrated a process by which careful interpretations of item non-responses could elucidate students' successful use of and benefit from accommodations.

The final study on implementation of accommodations (Lee et al.) addressed implementation practices and also yielded information about accommodations use patterns. While Lee and colleagues included universal tools, such as line-reader and highlighter, and designated supports, such as masking, we focus our review of findings on accommodations practices and use patterns. Similarly, although Lee and colleagues reported on findings for several student groups in the grade 6 population, we focus our discussion on students with disabilities and English learners with disabilities. Eligibility for using accommodations was similar across statewide English language arts (ELA) and mathematics assessments, with some differences between student groups. Some students with disabilities and English learners with disabilities did not use accommodations that they were eligible to use. About 83% of eligible grade 6 students with disabilities and about 86% of eligible grade 6 English learners with disabilities who were provided text-to-speech on ELA assessments were recorded as using that accommodation throughout the entire assessment. In contrast, partial text-to-speech—limited to designated portions of the assessment like directions or item stimuli—was only actually used by about 15% of students with disabilities and 17% of English learners with disabilities who were provided that version of the accommodation. Only about 29% of eligible students with disabilities and 25% of English learners with disabilities who were eligible to take a paper and pencil version of the assessment actually used it on ELA tests, and 10% of eligible students with disabilities and 15% of eligible English learners with disabilities actually used it on math tests. Lee and colleagues concluded that under-use of accommodations can occur.

## Discussion

---

This series of NCEO accommodations research reports has evolved since its inception in 1999 (Thompson et al., 2002). The report was originally published every two to three years, but became an annual publication with the 2017 report (Rogers et al., 2020). Since 2017, the annual report has identified a rather consistent trend of 10 to 12 relevant studies each year for over five years (Rogers et al., 2020, 2021; Rogers, Thurlow, et al., 2022; Rogers, Ressa, et al., 2022).

We identified three themes across the 2021 accommodations research reviewed. First, three studies presented evidence that highlighted the importance of considering student perceptions and preferences in relation to accommodations. Second, most of the research studies addressed mathematics or other STEM topics. Finally, more of research on assessment accommodations in 2021 considered paper and pencil administration as an accommodation in an otherwise digital environment than in previous years.

The topic of student perceptions is often addressed in accommodations research. However, this set of studies emphasized the importance of student perceptions in identifying appropriate assessment and classroom accommodations. Davis and colleagues compared the performance of high school students on mathematics assessments using three response options: paper and pencil, digital stylus, and keyboard. The student survey revealed that although students performed similarly using all three response options, they preferred the paper-and-pencil option. Long and colleagues also studied the effects of accommodations on student performance in mathematics. Students were asked to use two types of algebra tiles: concrete and virtual. As in the Davis and colleagues study, students performed similarly under both conditions. In both studies, students' performance was not significantly impacted by the different accommodations, but the students did express preferences. These studies together suggest a need for educators to consider student preferences when selecting accommodations.

We also observed students' individual preferences in another mathematics study. Zhang and Rivera conducted a mixed methods inquiry of students' problem-solving experiences and found that using a standard set of accommodations for all students with math difficulties was not sufficient. Instead, they concluded that a standard protocol of accommodations may not sufficiently meet students' individual needs and suggested that more individualized accommodations are necessary.

There has been an increase in the proportion of accommodations research associated with mathematics from 50% in 2019 to 80% in 2021. In 2021, accommodations encompassed high and low technology supports, with students in some studies indicating preferences for low tech supports (Davis et al.; Long et al.).

An emerging trend in accommodations research is the treatment of paper-and-pencil as an accommodation with computer-based assessments being the norm. In their report on patterns of accommodations use on grade 6 statewide assessments, Lee and colleagues included the option to have test items printed on paper as an accommodation. Davis and colleagues found that high school students preferred the paper-and-pencil response option to the use of a digital stylus or keyboard. Wang and colleagues compared student performance on the ACT when completed online or in the accommodated condition of paper-and-pencil format. Finally, Marble-Flint and colleagues compared kindergarten students' reading comprehension when read to from an electronic tablet to their comprehension when read to from a traditional storybook. As we continue this series of annual reports, we anticipate continuing to see an increase in the number of studies that consider the digital environment as the norm, with the traditional paper-and-pencil format being considered an accommodation.

## Report References

---

(References in the report to studies that were part of the 2021 accommodations research analysis are not included in this list. They are in the separate list titled 2021 K–12 Accommodation References.)

Blank, M., Rose, S. A., & Berlin, L. J. (2003). *Preschool language assessment instrument (2nd ed.; PLAI-2)* (TC022912) [Database record]. ETS TestLink.

Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged instrument. *Psychological Reports*, 99(1), 166–170. <https://doi.org/10.2466/pr0.99.1.166-170>

Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. Nova Science Publishers.

Connolly, A. J. (2007). *Key math3 diagnostic assessment (KeyMath-3 DA)* (TC024100) [Database record]. ETS TestLink.

Coombs, A., DeLuca, C., LaPointe-McEwan, D., & Chalas, A. (2018). Changing approaches to classroom assessment: An empirical study across teacher career stages. *Teaching and Teacher Education*, 71, 134–144. <https://doi.org/10.1016/j.tate.2017.12.010>

Cormier, D. C., Altman, J., Shyyan, V., & Thurlow, M. L. (2010). *A summary of the research on the effects of test accommodations: 2007–2008* (Technical Report 56). National Center on Educational Outcomes.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test (4th ed.; PPVT-4)*. Pearson Assessments.

Harper Collins Publishing. (2019, December). *I can read! Levels*. <https://www.icanread.com/levels/>

Johnstone, C. J., Altman, J., Thurlow, M. L., & Thompson, S. J. (2006). *A summary of the research on the effects of test accommodations: 2002 through 2004* (Technical Report 45). National Center on Educational Outcomes.

Maher, C. A., & Yankelewitz, D. (Eds.). (2017). *Children's reasoning while building fraction ideas*. Springer.

Rogers, C. M., Christian, E. M., & Thurlow, M. L. (2012). *A summary of the research on the effects of test accommodations: 2009–2010* (Technical Report 65). National Center on Educational Outcomes.



Rogers, C. M., Lazarus, S. S., & Thurlow, M. L. (2014). *A summary of the research on the effects of test accommodations, 2011–2012* (Synthesis Report 94). National Center on Educational Outcomes.

Rogers, C. M., Lazarus, S. S., & Thurlow, M. L. (2016). *A summary of the research on the effects of test accommodations: 2013–2014* (NCEO Report 402). National Center on Educational Outcomes.

Rogers, C. M., Lazarus, S. S., & Thurlow, M. L. (2021). *A summary of the research on the effects of K–12 test accommodations: 2018* (NCEO Report 423). National Center on Educational Outcomes.

Rogers, C. M., Lazarus, S. S., Thurlow, M. L., & Liu, K. K. (2020). *A summary of the research on the effects of K–12 test accommodations: 2017* (NCEO Report 418). National Center on Educational Outcomes.

Rogers, C. M., Ressa, V. A., Thurlow, M. L., & Lazarus, S. S. (2022). *A summary of the research on the effects of K–12 test accommodations: 2020* (NCEO Report 436). National Center on Educational Outcomes.

Rogers, C. M., Thurlow, M. L., Lazarus, S. S., & Liu, K. K. (2019). *A summary of the research on the effects of test accommodations: 2015–2016* (NCEO Report 412). National Center on Educational Outcomes.

Rogers, C. M., Thurlow, M. L., Lazarus, S. S., Ressa, V. A., & Swadek, G. S. (2022). *A summary of the research on the effects of K–12 test accommodations: 2019* (NCEO Report 433). National Center on Educational Outcomes.

Schulenberg, S. E., & Melton, A. (2008). The computer aversion, attitudes, and familiarity index (CAAFI): A validity study. *Computer in Human Behavior*, 24(6), 2620–2638. <https://doi.org/10.1016/j.chb.2008.03.002>

Schulenberg, S. E., Yutzenka, B. A., & Gohm, C. L. (2006). The computer aversion, attitudes, and familiarity index (CAAFI): A measure for the study of computer-related constructs. *Journal of Educational Computing Research*, 34(2), 129–146. <https://doi.org/10.2190/45B4-GMH7-GEQB-T1H1>

Shao, C., Li, J., & Cheng, Y. (2016). Detection of test speededness using change-point analysis. *Psychometrika*, 81(4), 1118–1141. <https://doi.org/10.1007/s11336-015-9476-7>

Smarter Balanced Assessment Consortium. (2016, January). *Smarter Balanced Assessment Consortium accommodations and accessibility framework*. <https://portal.smarterbalanced.org/library/en/accessibility-and-accommodations-framework.pdf>

Smarter Balanced Assessment Consortium. (2017). *Smarter Balanced Assessment Consortium: Summative technical report*. <https://portal.smarterbalanced.org/library/en/2017-18-summative-assessment-technical-report.pdf>

Thompson, S., Blount, A., & Thurlow, M. (2002). *A summary of the research on the effects of test accommodations: 1999 through 2001* (Technical Report 34). National Center on Educational Outcomes.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children (5th ed.; WISC-V)*. Pearson.

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence (3rd ed.; WPPSI-III)* (TC022144) [Database record]. ETS TestLink.

Zenisky, A. L., & Sireci, S. G. (2007). *A summary of the research on the effects of test accommodations: 2005–2006* (Technical Report 47). National Center on Educational Outcomes.

## 2021 K–12 Accommodation References (N=12)

---

Bohack, K. (2021). *The effect of computer familiarity on computer-based assessment in mathematics* (Publication No. 28317900) [Doctoral dissertation, Hofstra University]. ProQuest Dissertations and Theses Global.

Caldone, D. R. (2021). *The tensions of teacher assessment identity: What secondary teachers think, do, and learn from experience* (Publication No. 28543654) [Doctoral dissertation, Robert Morris University]. ProQuest Dissertations and Theses Global.

Davis, L. L., Morrison, K., Schnieders, J. Z.-Y., & Marsh, B. (2021). Developing authentic digital math assessments. *Journal of Applied Testing Technology*, 22(1), 1–11. <http://www.jattjournal.net>

Gilbert, K., Kranzler, J., & Benson, N. (2021). An independent examination of the equivalence of the standard and digital administration formats of the Wechsler Intelligence Scale for Children-5th Edition. *Journal of School Psychology*, 85, 113–124. <https://doi.org/10.1016/j.jsp.2021.01.002>

Lee, D., Buzick, H., Sireci, S. G., Lee, M., & Laitusis, C. (2021). Embedded accommodation and accessibility support usage on a computer-based statewide achievement test. *Practical Assessment, Research, and Evaluation*, 26(1), article 25. <https://doi.org/https://doi.org/10.7275/21926674>

Long, H., Bouck, E., & Domka, A. (2021). Manipulating algebra: Comparing concrete and virtual algebra tiles for students with intellectual and developmental disabilities. *Exceptionality*, 29(3), 197–214. <https://doi.org/10.1080/09362835.2020.1850454>

Marble-Flint, K. J., Strattman, K. H., & Schommer-Aikins, M. A. (2021). Comprehension scores among young neurotypical children and children with autism: Paper and iPad® storybooks. *Journal of Special Education Technology*, 36(4), 311–320. <https://doi.org/10.1177/0162643420924197>

Nnoli, K. C. (2021). *A generic qualitative inquiry of special educators' experiences preparing students for high-stakes tests* (Publication No. 28722335) [Doctoral dissertation, Capella University]. ProQuest Dissertations and Theses Global.

Wang, S., Li, D., & Steedle, J. (2021). *An investigation of speededness as a possible explanation for mode effects on the ACT* (Technical Brief No. 2021-R2142). ACT. <https://www.act.org/content/dam/act/unsecured/documents/2021/R2142-speededness-mode-effects-2021-11.pdf>

Zebehazy, K. T., & Wilton, A. P. (2021). Graphic reading performance of students with visual impairments and its implication for instruction and assessment. *Journal of Visual Impairment & Blindness*, 115(3), 215–227. <https://doi.org/10.1177/0145482X211016918>

Zhang, D., Indyk, A., & Greenstein, S. (2021). Effects of schematic chunking on enhancing geometry performance in students with math difficulties and students at risk of math failure. *Learning Disability Quarterly*, 44(2), 82–95. <https://doi.org/10.1177/0731948720902400>

Zhang, D., & Rivera, F. D. (2021). Predetermined accommodations with a standardized testing protocol: Examining two accommodation supports for developing fraction thinking in students with mathematical difficulties. *Journal of Mathematical Behavior*, 62(100861). <https://doi.org/10.1016/j.jmathb.2021.100861>

## Appendix A

### Research Characteristics for K–12 Studies in 2021

<b>Authors</b>	<b>Publication Type</b>	<b>Research Type</b>	<b>Research Design</b>	<b>Data Collection Source</b>	<b>Collection Instrument</b>
Bohack	Dissertation	Quantitative	Quasi-experimental	Primary	Survey; Test
Caldone	Dissertation	Mixed	Descriptive Quantitative	Primary	Interview Protocol; Survey
Davis et al.	Journal Article	Mixed	Quasi-experimental	Primary	Survey; Test
Gilbert et al.	Journal Article	Quantitative	Quasi-experimental	Primary	Test
Lee et al.	Journal Article	Quantitative	Descriptive Quantitative	Secondary	Other
Long et al.	Journal Article	Mixed	Experimental	Primary	Interview Protocol; Observations; Test
Marble-Flint et al.	Journal Article	Quantitative	Experimental	Primary	Test
Nnoli	Dissertation	Qualitative	Descriptive Qualitative	Primary	Interview Protocol
Wang et al.	Report	Quantitative	Descriptive Quantitative	Primary	Survey; Test
Zebehazy & Wilton	Journal Article	Quantitative	Quasi-experimental	Primary	Survey; Test
Zhang et al.	Journal Article	Quantitative	Quasi-experimental	Primary	Grades; Test
Zhang & Rivera	Journal Article	Qualitative	Descriptive Qualitative	Primary	Test

## Appendix B

### Research Purposes for K–12 Studies in 2021

Authors	Stated Research Purpose	Effects (SwD)	Effects (without disabilities)	Effects (both)	Perceptions	Implementation/Use	Issues	Summarize	Test Items	Validity
Bohack	Compare student performance on paper-based tests to computer-based tests; report on the effects of mode-related factors, particularly familiarity with computer use, including differences in scores and the degree of difference.			P			X			
Caldone	Study the beliefs (self-perceptions and knowledge) of middle and high school teachers on assessing students in general, and providing accommodations in particular; describe how teachers' beliefs were linked to their practices on assessment and accommodations; discuss issues including appropriate reasons for and applications of accommodations; summarize related research literature.				P	X	X	X		
Davis et al.	Compare student performance on three item response formats on a mathematics assessment (paper-and-pencil handwriting, digital stylus, and keyboard input) and report on student preferences among the three item response formats.		P		X					
Gilbert et al.	Investigate the effects of digital and paper-based test formats on intelligence test composite and subscale scores for students without disabilities; compare construct equivalence and measurement unit equivalence for two test formats.		P							X
Lee et al.	Describe implementation practices regarding accessibility supports and accommodations across four student groups (students with disabilities, English learners, English learners with disabilities and general education students); report on state-wide and district-level analyses on over-use and under-use of accommodations.					P				

Authors	Stated Research Purpose	Effects (SwD)	Effects (without disabilities)	Effects (both)	Perceptions	Implementation/Use	Issues	Summarize	Test Items	Validity
Long et al.	Compare the effectiveness of two types of mathematics manipulatives (concrete algebra tiles and virtual algebra tiles) in improving accuracy, independence and speed of students with intellectual disabilities solving linear algebra problems; report on student perceptions and preferences of the two manipulative types.	P			X					
Marble-Flint et al.	Compare kindergarten students' reading comprehension of two student groups (students with ASD and neuro-typical peers) when presented with storybooks in three formats: paper with teacher read aloud, electronic tablet with teacher read aloud, and electronic tablet with built in read aloud.			P						
Nnoli	Report on the perceptions of special education teachers in preparing students for high-stakes tests; report on teacher implementation of accommodation practices; summarize research literature related to experiences of special education teachers.				P	X		X		
Wang et al.	Compare student performance and degrees of speededness on a nationally recognized achievement test (ACT) presented in paper and online testing formats; discuss the issue of speededness as a possible factor related to assessment performance; report on student perceptions of testing formats.			P	X		X			
Zebehazy & Wilton	Compare the performance of students with visual impairments on graphics-based tasks in two presentation formats (tactile graphics and print graphics, including enlarged print); report on student perceptions of difficulty between the two presentation formats; discuss the relationship between teacher-rated student characteristics and student performance on graphics-based tasks.	P			X		X			

Authors	Stated Research Purpose	Effects (SwD)	Effects (without disabilities)	Effects (both)	Perceptions	Implementation/Use	Issues	Summarize	Test Items	Validity
Zhang et al.	Compare the effects of presentation accommodations for high school students with math difficulties (both students with and without disabilities) including visual chunking, use of contrasting colors, and black and white format; investigate the effects of student use of reference guides with geometric theorems and postulates; investigate the moderating effects of problem difficulty level and students' general knowledge.			P		X			X	
Zhang & Rivera	Compare the effects of two accommodations to a non-accommodated condition on student achievement in mathematics: (a) with important words or phrases in bold font and additional explanatory information, and (b) warm-up example items at a less complex level (e.g., whole numbers rather than fractions); report on whether and to what extent accommodations met student needs; summarize related research literature.	P				X		X		



## Key for Appendix B

Effects [SwD]	Compare effects of accommodations on assessment scores [only students with disabilities]
Effects [non]	Compare effects of accommodations on assessment scores [only students without disabilities]
Effects [both]	Compare effects of accommodations on assessment scores [both students with and without disabilities]
Perceptions	Study/compare perceptions and preferences about use
Implementation/Use	Report on implementation practices and accommodations use
Issues	Discuss issues related to test accommodations
Summarize	Summarize research on test accommodations
Test Items	Compare test items across assessment formats
Validity	Investigate test validity under accommodated conditions
P	Primary purpose
X	Additional purpose

## Appendix C

### Instrument Characteristics for K–12 Studies in 2021

**Table C-1. Instrument Types and Specific Instruments Used, and Their Sources**

<b>Authors</b>	<b>Instrument Types and Descriptions</b>	<b>Count</b>
Bohack	<p><b>Author (survey):</b> The computer familiarity survey was developed by this dissertation researcher, to be different from an already-established survey: Computer Aversion, Attitudes, and Familiarity Index (CAAFI; Schulenberg et al., 2006; Schulenberg &amp; Melton, 2008). The survey had 10 items, on a rating scale of 1–7, with 4 being neutral and higher scores indicating less aversion, more positive attitude, and greater familiarity.</p> <p><b>Researcher (test):</b> Two forms of math test constructed (by author) from state assessment, with algebra items linked to state content standards.</p>	2
Caldone	<p><b>Author (interview protocol):</b> 11 interview items requesting details relevant to teachers' survey responses on assessment perceptions and practices.</p> <p><b>Researchers (surveys):</b> Teachers' Conception of Assessment (TCoA; Brown, 2016) and Approaches to Classroom Assessment Inventory (ACAI; Coombs et al., 2018) were combined into one teacher survey.</p>	2
Davis et al.	<p><b>Author (survey):</b> In response to this author-developed student survey, students described their experiences, perceptions, and preferences about the assessment response options.</p> <p><b>Researcher (test):</b> A mathematics assessment, with 18 constructed response items, created by ACT content developers specifically for this study, over content including algebra, geometry, and numbers/quantities. Half of the items called for math process skills such as traditional computation, and the other half required math practice skills in which students provided answers and explained their rationales, such as demonstrating the application of formulas to reach solutions.</p>	2
Gilbert et al.	<p><b>Norm-ref Ability:</b> Wechsler Intelligence Scale for Children, fifth edition (WISC-V; Wechsler, 2014), subtest scores and full-scale IQ (FSIQ) scores measured general cognitive ability.</p>	1
Lee et al.	<p><b>Other:</b> Smarter Balanced extant statewide dataset for ELA and math tests (Smarter Balanced Assessment Consortium, 2016, 2017); these did not include actual test scores, but rather, other data for the student participants.</p>	1

Authors	Instrument Types and Descriptions	Count
Long et al.	<p><b>Author (interview protocol):</b> Four interview questions addressed social validity, inquiring from students' their experiences with the versions of manipulatives.</p> <p><b>Author (observations):</b> Students' task completion time was documented, and the degree of task support/independence was rated based on the system of least prompts—identifying whether prompts, and which level of prompts, were used.</p> <p><b>Researcher (test):</b> Pre-test of algebra skills performance of less than 20%; set of two-step addition or subtraction linear algebra problems.</p> <p><b>Norm-ref Ach:</b> The KeyMath3™ diagnostic assessment (KeyMath-3; Connolly, 2007) provided grade equivalence scores, separate from participants' task performances.</p> <p><b>Norm-ref Ability:</b> Wechsler Preschool and Primary Scale of Intelligence-3rd edition (WPPSI-III; Wechsler, 2002) and Wechsler Intelligence Scale for Children-fifth edition (WISC-V; Wechsler, 2014) provided general cognitive ability—through full-scale scores, separate from participants' task performances.</p> <p><b>Other (screening):</b> For screening, teacher-reported proficiency with one-step, single-digit mathematics operation tasks.</p>	5
Marble-Flint et al.	<p><b>Researcher (test):</b> Scoring rubric and screener with Level 1 and Level 2 questions, developed from the Preschool Language Assessment Inventory-second edition (PLAI-2; Blank et al., 2003). Reading texts at the Shared Reading level, linked to the I Can Read™ storybook series (HarperCollins, 2019).</p> <p><b>Norm-ref Ability:</b> The Peabody Picture Vocabulary Test, 4th Edition (PPVT-4, Dunn &amp; Dunn, 2007) provided aptitude scores on language competence and potential impairments; the researchers matched pairs of participants according to PPVT-4 scores to ensure equivalent groups.</p>	2
Nnoli	<p><b>Author (interview protocol):</b> A set of 12 guiding questions were asked in each interview, eliciting special education teachers' experiences when preparing students to participate in tests, including their feelings, helpful actions, challenges and how they manage them, and their reflections on themselves currently as well as early in their careers.</p>	1
Wang et al.	<p><b>Researcher (survey):</b> Student perceptions were gathered from two items, with 5-point scaling, from the student survey that was administered to all ACT participants, specifically "Q8. Please provide your level of agreement with the following statements: I had enough time to finish the English/mathematics/reading/science test" and "Q17. How helpful or unhelpful was the [on-screen] timer for managing your pace throughout the test?"</p> <p><b>Crit-ref Ach:</b> ACT in English, math, reading, science extant nationwide dataset. Three data sets were gathered: in October 2019, December 2019, and February 2020 utilizing the 2019–2020 school year version of ACT.</p> <p><b>Other:</b> Speededness was detected using the change-point analysis procedure (Shao et al., 2016).</p>	3

<b>Authors</b>	<b>Instrument Types and Descriptions</b>	<b>Count</b>
Zebehazy & Wilton	<p><b>Author (survey):</b> Students rated the difficulty levels (1=very easy; 5=very hard) of the tasks when using different graphic formats.</p> <p><b>Researcher (test):</b> A test-like task assessing science, technology, engineering, arts, and mathematics (STEAM) content comprised one to three multiple-choice questions for each of the five graphic types, which included: (a) bar graph, (b) Venn diagram, (c) shape rotation, (d) geometry area problem, and (e) map. Questions were leveled by schooling level (elementary, middle, and high school) but all used the same graphics.</p> <p><b>Other:</b> Teacher-reported student data were also analyzed for correlations of these various factors in relation to performance. Demographic, academic, and disability information were reported by participants' teachers—including teachers' own impressions of students—such as student history of instruction on graphics, IEP goals, and ability level in content areas.</p>	3
Zhang et al.	<p><b>Researcher (test):</b> A set of 22 mathematics problems, addressing geometry content—such as calculation of parallel lines, description of spatial relations, application of theorems or postulates—and incorporating three different difficulty levels: simple one-step, difficult one-step, and multi-step.</p> <p><b>State Test:</b> For screening, potential participants' failure on an unidentified state's mathematics assessment.</p> <p><b>Grades:</b> For screening, potential participants' course grade average of C or below in a high school geometry class.</p>	3
Zhang & Rivera	<p><b>Researcher (test):</b> Mathematics, specifically fraction sense, was the academic content tested by a set of five test-like tasks. Tasks included number line representations to story problems, concrete manipulative representations, and part-of-a-whole relationships, and were equivalent in difficulty across test versions. Some tasks were drawn, and then modified, from Maher and Yankelewitz (2017).</p>	1

### Key for Appendix Table C-1

<b>Instrument Type</b>	<b>Type Abbreviations</b>	<b>Number of Studies</b>
Surveys or academic tests developed by professionals or researchers through work outside of current study	Researchers (survey, test)	9
Non-academic protocols or surveys developed by study author/s	Author (survey/ interview/ observation)	6
Norm-referenced academic achievement measures	Norm-ref Ach	1
Norm-referenced cognitive ability measures	Norm-ref Ability	3
Course grades	Grades	1
Criterion-referenced academic achievement measures	Crit-ref Ach	1
Norm-referenced academic achievement measures	Norm-ref Ach	1
State criterion-referenced assessments	State Test	1
Other	Other	4

**Table C-2. Content Areas Assessed**

<b>Authors</b>	<b>Math</b>	<b>Reading</b>	<b>Writing</b>	<b>Other LA</b>	<b>Science</b>	<b>Social Studies</b>	<b>Cognitive Skills</b>	<b>N</b>
Bohack	•							1
Caldone								0
Davis et al.	•							1
Gilbert et al.							•	1
Lee et al.	•			•				2
Long et al.	•							1
Marble-Flint et al.		•						1
Nnoli								0
Wang et al.	•	•		•	•			4
Zebehazy & Wilton	•				•			2
Zhang et al.	•							1
Zhang et al.	•							1
<b>TOTAL</b>	<b>8</b>	<b>2</b>	<b>0</b>	<b>2</b>	<b>2</b>	<b>0</b>	<b>1</b>	

## Appendix D

### Participant Characteristics for K–12 Studies in 2021

<b>Authors</b>	<b>Unit of Analysis</b>	<b>Sample Size</b>	<b>Percent of Student Sample with Disabilities</b>	<b>School Level</b>	<b>Disability Categories Included in Sample</b>
Bohack	Students	53	19%	Grades 9–11 / high school	Not specified
Caldone	Educators	131	N/A	N/A	N/A
Davis et al.	Students	100	0%	High school	None
Gilbert et al.	Students	65	0%	Grades K–12 / elementary, middle and high school	None
Lee et al.	Students	481,282	14%	Grade 6 / middle school	Not specified; None
Long et al.	Students	3	100%	Grade 8 / middle school	ID
Marble-Flint et al.	Students	30	50%	Children aged 4–6 / kindergarten	A; None
Nnoli	Educators	8	N/A	N/A	N/A
Wang et al.	Students	16,580	Unknown <sup>a</sup>	Grade 12 / high school	Not specified
Zebehazy & Wilton	Students	40	100%	Grades 4–12 / elementary, middle, and high school	VI
Zhang et al.	Students	33	52%	Grade 9 / high school	LD; AP
Zhang & Rivera	Students	23	100%	Grade 6–8 / middle school	LD; AP; S/L

<sup>a</sup> Unknown: Wang et al. did not specify the numbers of students with and without disabilities, and did not use comparison groups of student with and without disabilities.

**Key for Appendix D**

A	Autism
AP	Attention problem
ID	Intellectual disability
LD	Learning disability
S/L	Speech/Language impairment
VI	Visual impairment/blindness
None	Students without disabilities
Not specified	Students with disabilities, no categories reported
N/A	No student participants

## Appendix E

---

### Accommodations Studied for K–12 Studies in 2021

**Table E-1. All Accommodations by Study**

<b>Authors</b>	<b>Accommodations</b>
Bohack	Bundle/aggregated set: Electronic administration with embedded accommodations including answer elimination, flag questions, notepad, review tool
Caldone	Not specified
Davis et al.	Digital stylus; Keyboarding
Gilbert et al.	Electronic administration (with digital tablet)
Lee et al.	Oral delivery (text-to-speech, partial or full); Highlighter (by student); Line reader; Masking
Long et al.	Concrete manipulatives; Virtual manipulatives
Marble-Flint et al.	Electronic administration (with digital tablet); Text-to-speech; Oral delivery (live in-person and pre-recorded)
Nnoli	Not specified
Wang et al.	Paper/pencil, in otherwise digital environment
Zebehazy & Wilton	Tactile graphics [Note: both test conditions included enlarged print/font]
Zhang et al.	Visual chunking to draw attention to important details; Reference guide with formulas, theorems, and postulates
Zhang & Rivera	Annotations—bolding important words/phrases and providing explanations; Warm-up practice items using whole numbers prior to test items on fractions



**Table E-2. Presentation Accommodations Itemized by Study**

<b>Authors</b>	<b>Electronic administration</b>	<b>Text-to-speech device/software</b>	<b>Enlargement (font on screen/ print on paper)</b>	<b>Masking</b>	<b>Oral delivery, live/in-person</b>	<b>Paper/pencil, in otherwise digital environment</b>	<b>Practice items</b>	<b>Tactile graphics</b>	<b>Recorded delivery, human voice</b>	<b>Visual chunking</b>	<b>N</b>
Bohack	• <sup>a</sup>										1
Gilbert et al.	•										1
Lee et al.		•		•							2
Marble-Flint et al.	•	•			•			•			4
Wang et al.						•					1
Zebehazy & Wilton			•					•			2
Zhang et al.										•	1
Zhang & Rivera							•				1
<b>TOTAL studies (of 8)</b>	<b>3</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>

<sup>a</sup> Part of an aggregated set (“bundle”) of accommodations

**Table E-3. Equipment and Materials Accommodations Itemized by Study**

Authors	Electronic administration	Bold font for important words/phrases <sup>a</sup>	Flag questions	Highlighter, by student	Line reader	Manipulatives: concrete, virtual	Notepad (digital note space)	Reference guide (math)	Review tool	N
Bohack	•		• <sup>b</sup>				• <sup>b</sup>		• <sup>b</sup>	4
Gilbert et al.	•									1
Lee et al.				•	•					2
Long et al.						•				1
Marble-Flint et al.	•									1
Zhang et al.								• <sup>c</sup>		1
Zhang & Rivera		•								
<b>TOTAL studies (of 7)</b>	<b>3</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	

<sup>a</sup> “Annotation” includes explanations of the words or phrases

<sup>b</sup> Part of an aggregated set (“bundle”) of accommodations

<sup>c</sup> Reference guide had formulas, theorems, and postulates for math items

**Table E-4. Response Accommodations Itemized by Study**

Authors	Electronic administration	Answer elimination	Digital stylus	Keyboard	N
Bohack	•	• <sup>a</sup>			2
Davis et al.			•	•	2
Gilbert et al.	•				1
Marble-Flint et al.	•				1
<b>TOTAL studies (of 4)</b>	<b>3</b>	<b>1</b>	<b>1</b>	<b>1</b>	

<sup>a</sup> Part of an aggregated set (“bundle”) of accommodations

**Table E-5. Scheduling and Timing Accommodations Itemized by Study**

<b>Authors</b>	<b>N/A</b>	<b>N</b>
N/A		<b>0</b>
<b>TOTAL studies (of 12)</b>	<b>0</b>	

**Table E-6. Setting Accommodations Itemized by Study**

<b>Authors</b>	<b>N/A</b>	<b>N</b>
N/A		<b>0</b>
<b>TOTAL studies (of 12)</b>	<b>0</b>	

## Appendix F

### Findings for K–12 Studies in 2021

Authors	Findings Statement	Effects	Perceptions	Implement/ Use	Test Items	Validity
Bohack	<p>The analysis showed no significant mean differences in scores between the computer-based test (CBT) and the paper-based test (PBT). Further, no patterns were found when examining placement level, gender, and other factors in general. Scores between formats were deemed comparable and the formats were interchangeable. No significant correlation was found between computer familiarity CBT scores, and no significant correlations were observed between computer familiarity and CBT scores, PBT scores, differences in scores, or prior achievement. Thus, degree of familiarity with computers was not a factor influencing performance. The implication was that implementing computer-based assessments was deemed appropriate, as there were no indications that this administration format was detrimental to student participants.</p>	X		X		
Caldone	<p>Teachers perceived that they have had relative ease in providing accommodations during classroom assessments, and indicated relative difficulty in ensuring students' access (through accommodations) to the state assessments. Survey respondents averaged relatively strong agreement (4=agree; 6=strongly agree) on applying various accommodations practices, including consistently providing accommodations across all assessment tasks (mean=4.97), incorporating accommodations in classroom comprehensive or end-of-course assessments (mean=4.96). However, respondents indicated less concurrence on other related practices, such as communicating with parents on the purpose of accommodations (mean=3.86), and considering student needs and accommodations in relation to policies (mean=3.15). Of the components of assessment purpose, assessment process, and assessment fairness, accommodations practices were only, and most closely, associated with fairness. Interview responses pertaining to practices such as differentiating assessments based on the needs of students were also associated with the importance of accommodations.</p>		X	X		

Authors	Findings Statement	Effects	Perceptions	Implement/ Use	Test Items	Validity
Davis et al.	Mean performance scores showed no statistically significant differences across item response input conditions—that is, traditional paper-and-pencil responding, and the two comparison conditions, keyboard and digital stylus on screen. When comparing scores by response conditions for math content areas and for math construct type, there were slight group mean differences yet none that reached statistical significance. Student surveys showed preferences among the response conditions: paper and pencil was the most highly preferred, with digital stylus preferred to keyboard.	X	X			
Gilbert	Researchers compared performance of 65 students without disabilities on both the digital and paper forms of the Wechsler Intelligence Scale for Children, fifth edition (WISC-V; Wechsler, 2014) concluding that the two forms were not equivalent because student performance on the digital format was higher than the standard paper format. Because the study was limited to students without disabilities it is difficult to apply findings to students with disabilities.					X

Authors	Findings Statement	Effects	Perceptions	Implement/ Use	Test Items	Validity
Lee et al.	<p>Eligibility for using accommodations differed somewhat for grade 6 students with disabilities who were and were not English learners, yet was similar across statewide ELA and mathematics assessments. The eligibility rates for using text-to-speech software for oral delivery of ELA text passages were about 30% of students with disabilities and 46% of English learners with disabilities. The eligibility rates for paper-and-pencil testing instead of electronic administration were about 1% for both groups. Incidence of use for students with disabilities and English learners with disabilities was comparatively lower than eligibility to use accommodations, yet ranged among individual accommodations and between test content. For text to speech, 83% of eligible students with disabilities and 86% of eligible English learners with disabilities actually used the accommodation. Partial text-to-speech—that is, for selected words or items—was actually used by about 15% of students with disabilities and 17% of English learners with disabilities. For the paper-and-pencil format, on ELA tests, actual use rates were 29% of students with disabilities and 25% of English learners with disabilities; on math tests, rates were 10% of students with disabilities and 15% of English learners with disabilities. Lower rates of eligible students actually using assessment supports was termed “underuse” when analyzing district-level data across the state. The researchers demonstrated the procedures for identifying districts with underuse rates, noting the importance of addressing these implementation concerns. (Note: Highlighter, line reader, and masking were not identified by state policy as accommodations.)</p>			X		
Long et al.	<p>Students with intellectual disabilities in grade 8 were equally independent, accurate, and timely when using concrete (physical) manipulatives and virtual manipulatives. Concrete and virtual manipulatives were equally effective in supporting participants’ performance on algebra test-like tasks. The three participants demonstrated similar benefits when applying the two types of manipulatives during grade-level math tasks over time. All three participants indicated that they liked using the manipulatives, with a slight preference for the concrete tiles.</p>	X	X			

Authors	Findings Statement	Effects	Perceptions	Implement/ Use	Test Items	Validity
Marble-Flint et al.	All participants—with and without autism—performed significantly better with either form of the electronic tablet presentation, including via digital tablet with in-person oral delivery and with embedded oral delivery of recorded human voice, compared to in-person oral delivery with the paper book format. There were no significant differences in participant performance between the two electronic table presentation formats.	X				
Nnoli	Special educator interviewees reported several accommodations-related factors in preparing students for high stakes assessments. Interviewees reported that knowledge of students' IEP accommodations was very important. They noted the need for providing instruction on and practice opportunities for accommodations. Special educators reported that they believed it was important to ensure students received accommodations in the classroom and during testing, and that accommodations were necessary for students to demonstrate their knowledge during assessments. Teachers reported preparing students to use various accommodations including extended time, speech-to-text, human reader, calculator tool, voice or immersion reader, and scribe for responses.		X	X		
Wang et al.	Larger proportions of students' scores indicated speededness on the ACT in reading and English tests in general, and during the alternative paper version more than the typical online test. Speeded score patterns were observed when students seemed to respond too quickly to items than would be needed for students to choose answers for correctness. For both the reading and English tests, students' mean scores were higher on the online testing than on the paper format. About 40–46% of students testing in both the online and paper formats indicated on surveys that they had enough time to complete the test at similar rates for reading, math, and science tests. Nearly all students (94%) who took the ACT online indicated that the on-screen timer was helpful for pacing. Students perceived the English test to be less speeded; 73% agreed they had enough time. However, researchers found that students were most often flagged for speededness on the English test than other content tests.	X	X			

Authors	Findings Statement	Effects	Perceptions	Implement/ Use	Test Items	Validity
Zebehazy & Wilton	Significant differences were found for students with visual impairments in grades 4–12 in performance across medium types: print graphics users scored significantly better than tactile graphics users overall. Significant relationships were reported for teacher-rated factors on students' overall performance and specific graph types. These factors were: (a) an IEP goal on graphics, (b) math ability level, (c) graphics engagement frequency, (d) problem-solving skills, and (e) independence level using graphics. Task completion took longer for tactile graphics users, which researchers recommended should be factored into instruction and assessment practices. Researchers concluded that students needed more instruction and practice using tactile graphics during assessments in order to benefit from this accommodation. Student performance generally aligned with their teachers' ratings of student ability (e.g., students rated as highly independent in problem solving performed well on most tasks). Researchers found that the difficulty level perceived by students did not match their actual performance.	X	X			
Zhang et al.	All student groups—that is, students with difficulties in geometry who (a) had low-average classroom performance, (b) failed math assessments, and (c) had disabilities—improved in performance with chunking representations. The researchers extended the concept of “chunk” from just a physical shape to a schematic with meaning in relation to geometric properties, theorems, and postulates. Two possible moderating effects were reported, problem complexity and student knowledge of geometry. In terms of problem complexity, the effects of chunking were not significant on one-step problems but were significant on more complex problems that required greater working memory load. Results showed that students with math difficulties demonstrated significant improvement with chunking representation and reference guide while students at risk of failure did not. The researchers asserted that the accommodations reduced students' cognitive load by providing cognitive scaffolding, a process that was shown to be specifically beneficial for students with limited math knowledge and skill. Chunking provided a first level of scaffolding helping with recognition while the reference guide provided a second level of scaffolding for retrieval from long-term memory.	X		X	X	



Authors	Findings Statement	Effects	Perceptions	Implement/ Use	Test Items	Validity
Zhang & Rivera	The researchers reported that their qualitative inquiry of students' task problem-solving experiences yielded complex data. Students with math-related disabilities communicated on their thinking processes while completing tasks. A series of conclusions was reached: (a) Assessment performance data on response correctness or accuracy may not fully illuminate effects of accommodations, due to the individual differences in which students with math disabilities experience obstacles to problem-solving during testing. (b) A standard protocol of accommodations for all test takers with math disabilities may not be sufficiently useful in addressing these individual differences in needs during problem solving; instead, individualized accommodations could be needed. (c) Student participants' non-responses on the same items in different test conditions—with and without item annotations—require careful consideration of their meaning. The researchers suggested that students did not understand expectations despite simplified instructions; in other words, the annotations did not remove decoding-based obstacles. (d) Using whole number practice items prior to completing fraction items could support development of math reasoning capacity—that is, new learning.	X		X		
<b>TOTAL</b>		<b>8</b>	<b>6</b>	<b>5</b>	<b>1</b>	<b>1</b>

**Key for Appendix F**

Effects	Compare effects of accommodations on assessment scores
Implement/Use	Report on implementation practices and accommodations use
Perceptions	Study/compare perceptions and preferences about use
Test Items	Compare test items across assessment formats
Validity	Investigate test validity under accommodated conditions

INSTITUTE *on* COMMUNITY INTEGRATION

---

UNIVERSITY OF MINNESOTA

*NCEO is an affiliated center of the Institute on Community Integration*