

**A Summary of the
Research on the
Effects of K-12 Test
Accommodations: 2019**

NCEO Report 433



NCEO

National Center on
Educational Outcome:

**A Summary of the Research on the Effects
of K–12 Test Accommodations: 2019**

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

May 2022

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

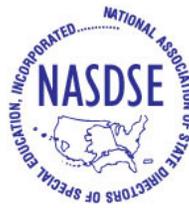
Rogers, C. M., Thurlow, M. L., Lazarus, S. S., Ressa, V. A., & Swadek, G. S. (2022). *A summary of the research on the effects of K-12 test accommodations: 2019* (NCEO Report 433). National Center on Educational Outcomes.

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
Terri Vandercook
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

Academic inquiry into the provision of accommodations during assessments has proven to have continued value for students with disabilities and for states setting accommodations policies. Key matters under ongoing investigation include the relative impacts of different types and forms of accommodations on student performance, educators' and students' knowledge and perceptions of accommodations, and broad patterns of selection and implementation of accommodations; findings from these investigations provide valuable information. Emerging issues across recent years include how large-scale testing delivered online through various platforms and devices has influenced the availability of accommodations, and related considerations of needed practices in applying new technologies to accessing tests for students with disabilities.

The purpose of this report is to provide an update on the state of the research on testing accommodations. Previous reports by the National Center on Educational Outcomes (NCEO) have covered research published since 1999. In this report, we summarize the research published in 2019. During 2019, 11 research studies were published on the topic of testing accommodations in the U.S. elementary and secondary education system.

Purpose of research: About four-fifths of the research published in 2019 converged around two purposes: the first was the comparison of the effects of accommodations on assessment scores for K–12 students with disabilities and the second was to report on implementation practices and accommodations use. The comparison of effects and implementation practices comprised 46% and 36% of the research respectively. The remainder of the studies—about one-fifth each—were either investigations of perceptions of accommodations or discussions of issues related to test accommodations. More than half of the studies had one purpose.

Research design: Over three-fifths of studies reported primary data by the researchers and less than two-fifths of studies used secondary (extant) data sets. Experimental or quasi-experimental designs were used in over half of the studies. Longitudinal and descriptive quantitative designs constituted about one-fifth of studies. One study implemented a correlation/prediction design. Researchers applied several quantitative and qualitative methodologies, specifically: tests, surveys, and observations.

Types of assessments, content areas: A variety of instruments was used across the studies. Over three-fifths of the studies used non-academic protocols or surveys developed by the study authors. Close to two-fifths of the studies used state criterion-referenced assessments. A few studies used survey or academic tests developed outside the study. Norm-referenced academic achievement measures and norm-referenced cognitive measures were each used in less than one-fifth of studies. One study used a criterion-referenced academic achievement test.

Participants: The majority of participants across the studies were K–12 students. Less than half of the studies had participants from more than one school level, usually middle school students

in combination with high school or elementary school students. Educators were respondents in two studies. Sample sizes had large variation; one study had three participants, while most had between 37 and 240 participants, and three studies with extant data sets each had over 10,000 participants.

Disability categories: Participants' disability categories varied with no category as the most represented. Participants with learning disabilities were represented in a little over one quarter of studies, with over 82 participants. Less than one-fifth of studies had a mix of disability categories including emotional behavioral disorders, learning disabilities, physical disabilities, speech/language impairments, and attention problems. The disability categories of visual impairments/blindness and traumatic brain injury were each investigated in one study. Three studies did not specify the disability categories of participants, and one study included no students with a designated disability category.

Accommodations: Presentation accommodations were the most frequently studied category of accommodations. Oral delivery, on a synchronous ("live") basis by test proctors or administrators, was the most-studied individual accommodation. A relatively large body of the studies published in 2019 comprised single-study investigations of specific accommodations (e.g., student reads aloud to self, dictated response to a speech recognition system).

Findings: Six studies analyzed the effects of accommodations. Half of these studies used technology as part of the accommodations, including electronic tablet, a multi-sensory tablet, and speech-to-text software. One-third of the studies included human proctors for prompting students, and one-third included oral delivery (self or human reader). Two-thirds of these studies showed that students performed better when provided with accommodations, in various content areas (reading, math, and writing). One-third of studies showed that using different modes of assessment (such as a multisensory tablet or electronic administration) resulted in performance similar to the traditional mode of assessment.

Over one-third of studies from 2019 reported findings on accommodations use and implementation practices. Patterns of accommodations use across populations of students were described in half of these studies, while educators' knowledge of accommodations and related implementation practices were detailed by researchers of the other half of these studies. Teachers' or students' perceptions of accommodations were relayed in only a couple studies. Procedures and approaches for engaging in test item comparison across assessment formats were explicated in one study.

Table of Contents

Executive Summary	iii
Overview.....	1
Review Process	1
Results.....	3
Publication Type	3
Purposes of the Research	4
Research Type and Data Collection Source.....	6
Data Collection Methods and Instruments.....	7
Content Area Assessed.....	10
Research Participants	11
School Level	12
Disability Categories.....	13
Types of Accommodations.....	14
Research Findings.....	16
Impact of Accommodations	17
Use and Implementation of Accommodations.....	18
Perceptions about Accommodations	19
Accommodations by Academic Content Assessments.....	20
Discussion	23
Report References.....	25
2019 K–12 Accommodation References (N=11).....	27
Appendix A: Research Characteristics for K–12 Studies in 2019	29
Appendix B: Research Purposes for K–12 Studies in 2019	30
Appendix C: Instrument Characteristics for K–12 Studies in 2019	33
Appendix D: Participant and Sample Characteristics for K–12 Studies in 2019	37
Appendix E: Accommodations Studied for K–12 Studies in 2019.....	38
Appendix F: Findings for K–12 Studies in 2019.....	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) of 2004 and by the 2015 reauthorization of 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), and 2018 (Rogers et al., 2021). The report summarizing the 2017 empirical studies narrowed the focus to K–12 research in the United States context.

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 2019. The academic literature described here encompasses empirical studies of performance comparability, as well as investigations into accommodations use, implementation practices, and perceptions of the 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 past accommodations research syntheses, a number of sources were accessed to complete the review of the K–12 accommodations research published in 2019. Specifically, five research databases were consulted: Educational Resources Information Center (ERIC), PsycINFO, Academic Search Premier, Digital Dissertations, and Educational Abstracts. To help confirm the thoroughness of our searching, we used the Web search engine Google Scholar to locate additional research. 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 were also searched for relevant publications. These organizations included Behavioral Research and Teaching (BRT) at the University of Oregon (<https://www.brtprojects.org/publications/>), the College Board Research Library (<http://research.collegeboard.org>), the National Center for Research on Evaluation, Standards, and Student Testing (CRESST; <http://cresst.org/education/>), and the Wisconsin Center for Educational Research (WCER; <https://www.wcer.wisc.edu/publications>).

The initial search was completed in December, 2019. A second search was completed in March, 2020, to ensure that all articles published in 2019 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 review using several criteria.

1. This analysis included only research published or defended (in doctoral dissertations) in 2019.
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 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 was focused on research on students in schools in the United States; 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 previous NCEO accommodations research reports).

These limitations 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>). This Bibliography will continue to include research in non-U.S. settings. Also, 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 accommodations research in the K–12 system that was published in 2019, the studies are examined and summarized on the following features: (a) publication type; (b) purposes of research; (c) research type and data collection source; (d) assessment or data collection focus; (e) characteristics of the independent and dependent variables under study; and (f) comparability of findings between studies in similar domains.

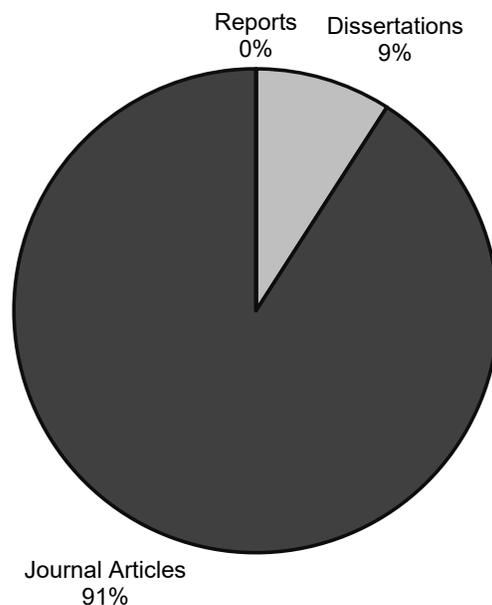
Results

Publication Type

Eleven studies were published between January 2019 and December 2019. As shown in Figure 1, 10 of the 11 studies were journal articles, and one was a dissertation; zero were published

professional reports published by research organizations or entities (e.g., BRT at the University of Oregon). The total number of studies published on accommodations in the K–12 educational context in 2019 ($N=11$) was the same as in 2018. When we apply our current more stringent inclusion criteria to NCEO’s report summarizing accommodations research in 2015–2016 (Rogers et al., 2020), the resulting analyses indicated that the number of published K–12 studies was nine in 2015 and nine in 2016. The numbers of journal articles increased, from four in 2015, to seven in 2016, to seven in 2017, to nine in 2018, to 10 in 2019. The largest variability in type was for dissertations, with five in 2015, two in 2016, seven in 2017, two in 2018, and one in 2019. The absence of reports from research organizations has continued, with zero in 2015, 2016, 2017, 2018, and 2019. This (2019) review included 10 journal articles from nine journals, including two articles from one academic journal. In 2018, there were nine journal articles from seven journals (Rogers et al., 2021); in 2017 there were seven journal articles from seven journals (Rogers et al., 2020). In 2016 there were seven articles from seven journals, and in 2015 there were four articles from four journals (Rogers et al., 2019). Appendix A presents information about the publication type of each study.

Figure 1. Percentage of Accommodations Studies by Publication Type in 2019



Purposes of the Research

The K–12 accommodations research published in 2019 had several purposes. The primary purposes of the 11 studies included in this review are shown in Table 1. Six studies indicated a single purpose (see Appendix B). The remaining five studies described multiple purposes.

In these cases, we identified the “primary purpose” based on the narrative description of the research questions, title of the work, or the first-mentioned purpose in the text.

The most frequent primary purpose for research published during 2019 was to compare scores to identify the effects of accommodations on test performance. Of these studies, four focused on the scores of only students with disabilities and one focused on only students without disabilities. Also significant in this review, three studies reported on the implementation and use of accommodations. The balance of the studies focused on comparing perceptions of accommodations, comparing test items across formats, and investigating test validity under accommodated conditions.

Table 1. Primary Purpose of K–12 Studies in 2019

Purpose	Number of Studies	Percent of Studies
Compare scores	5	45%
only students with disabilities (4 studies; 36% of studies)		
only students without disabilities (1 study; 9% of studies)		
both students with and without disabilities (0 studies; 0% of studies)		
Report on implementation practices and accommodations use	3	27%
Study/compare perceptions and preferences about use	1	9%
Compare test items across assessment formats	1	9%
Investigate test validity under accommodated conditions	1	9%
Discuss issues related to test accommodations	0	0%
Summarize research on test accommodations	0	0%

Although six studies identified single purposes, five studies had multiple purposes. Table 2 shows the primary, secondary and, in one case, tertiary purposes of the 11 studies. Out of all studies’ identified purposes, the most frequent was to analyze the effects of accommodations through comparing performance data ($n=6$); this includes comparing scores of only students with disabilities ($n=5$) and only students without disabilities ($n=1$). The next most frequent research purpose was reporting on the implementation and use of accommodations by four, or 36%, of the studies. The balance of the study purposes included comparing perceptions and preferences about use of accommodations, comparing test items, investigating test validity, discussing issues related to test accommodations, and summarizing research on test accommodations.

Table 2. All Purposes of K–12 Studies in 2019

Purpose	Number of Studies	Percent of Studies
Compare scores	6	55%
only students with disabilities (5 studies; 45% of studies)		
only students without disabilities (1 study; 9% of studies)		
both students with and without disabilities (0 studies; 0% of studies)		
Report on implementation practices and accommodations use	4	36%
Study/compare perceptions and preferences about use	2	18%
Compare test items across assessment formats	1	9%
Investigate test validity under accommodated conditions	1	9%
Discuss issues related to test accommodations	2	18%
Summarize research on test accommodations	1	9%

Note. Five studies (46%) had more than one purpose; therefore, numbers total more than the 11 studies represented, and percents total more than 100.

Appendix B presents an overview of the research purposes for the 11 studies in this review. The number of studies with a single purpose ($n=6$) was not a large difference from the number of studies with multiple purposes ($n=5$). One study (McCormack, 2019) had three identified research purposes: reporting on the implementation and use of accommodations, discussing issues related to test accommodations, and summarizing research on test accommodations. The majority of studies ($n=6$; 55%) investigated the impact of accommodations on student performance, either with or without disabilities. One study (Quesen & Lane, 2019) completed analyses of differential item functioning (DIF), using four different approaches, of item response patterns on a state mathematics assessment of students with and without disabilities.

Research Type and Data Collection Source

Experimental and quasi-experimental designs were the two most frequent types of accommodations research published in 2019, together comprising more than one-half of the 11 K–12 studies. Two studies employed descriptive quantitative designs, and researchers in two studies analyzed secondary data sets longitudinally. Only one study used correlational or predictive data. None of the studies used qualitative data; therefore, that design is not included in Table 3.

The researchers for most studies from 2019 gathered the data themselves (i.e., used primary source data); four (36%) of the studies relied on data from secondary sources. A similar proportion—more than 20 percent—of the studies we reviewed from 2017 drew data from secondary sources (Rogers et al., 2020). In contrast, in NCEO’s most recent previous report (Rogers et al., 2021), all 11 studies examined from 2018 had primary-sourced data, and none used only secondary data sources. (Appendix A presents research designs and data collection sources for individual studies.)

Table 3. Research Type and Data Collection Source for K–12 Studies in 2019

Research Type	Primary Source	Secondary Source	Total
Experimental	3	0	3
Quasi-experimental	2	1	3
Descriptive quantitative	2	0	2
Longitudinal	0	2	2
Correlation/Prediction	0	1	1
Totals	7	4	11

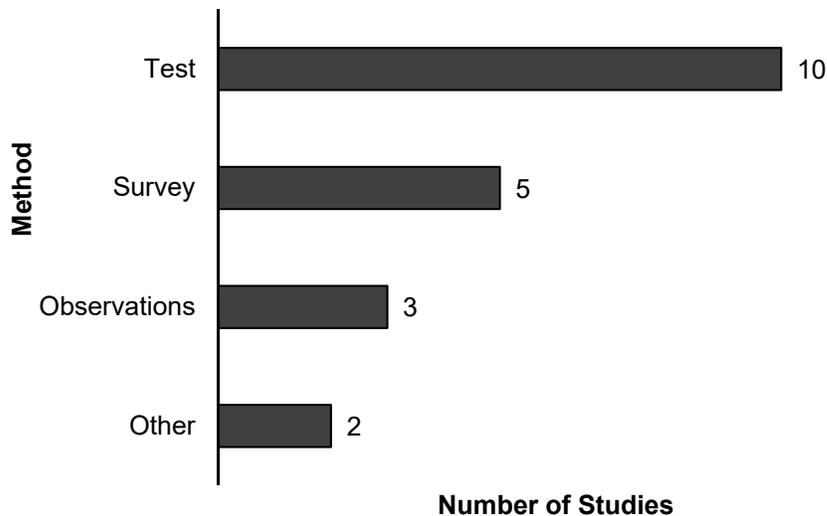
Data Collection Methods and Instruments

The 2019 research employed the methods shown in Figure 2 to collect study data. Nearly all of the studies ($n=10$) used performance data acquired through academic content testing. In some studies (e.g., Noakes et al., 2019), tests were administered as part of the study, while in others (e.g., Quesen & Lane, 2019), extant academic data sources were used. Surveys supplied data for nearly half of the 11 studies, including teacher surveys, student surveys, and parent surveys. Three studies (Carey et al., 2019; Castro et al., 2019; McCormack, 2019) engaged teachers as survey respondents, one study (Hahn et al., 2019) drew survey data from students, and one study (Kern et al., 2019) gathered demographic data about student participants from parents of participants. Observations were made in a plurality (27%) of studies in 2019. In contrast, a majority (55%) of the studies analyzed from 2018 employed observation (Rogers et al., 2021). Seven studies (64%) reported using more than one method or tool to gather data. The most common combination of collection methods was tests and surveys ($n=4$, 37%), and another frequent combination was tests and observations ($n=3$, 27%). See Appendix A for additional details about each study’s data collection methods.

All of the studies published in 2019 used some type of data collection instrument (see Table 4). The terms used in Table 4 are defined as follows:

- “Surveys” refers to items of an attitudinal or self-report nature.
- “Tests” is defined as course- or classroom-based.
- “Assessments” indicates statewide or large-scale assessments in scope.
- “Protocols” refers to sets of procedures, including observational.
- “Measures” refers to norm-referenced academic achievement or cognitive ability instruments.

Figure 2. Data Collection Methods Used in K–12 Studies in 2019



Note. Of the 11 studies reviewed for this report, five reported using two data collection methods and two reported using three data collection methods. Thus, the number of methods in this figure totals more than 11.

“Other” data collection methods are specified in Appendix C, Table C-1.

All of the instruments were placed into seven categories:

- Non-academic protocols or surveys developed by study authors
- Surveys or academic tests developed by education professionals or researchers using sources outside of current studies
- State criterion-referenced academic assessments
- Norm-referenced academic achievement measures
- Norm-referenced cognitive ability measures
- Non-state criterion-referenced academic achievement measures
- Other

Table 4. Data Collection Instrument Types for K–12 Studies in 2019

Instrument Type	Number of Studies^b	Percent of Studies^b
Non-academic protocols or surveys developed by study author/s	7	64%
State criterion-referenced academic assessments	4	36%
Surveys or academic tests developed by professionals or researchers using sources outside of current study	3	27%
Norm-referenced academic achievement measures	2	18%
Norm-referenced cognitive ability measures	2	18%
Non-state criterion-referenced academic achievement measures	1	9%
Other ^a	2	18%

^a Other: see Appendix C, Table C-1 for specific information in Kern et al., 2019; and Wise et al., 2019.

^b Seven studies (64%) used more than one type of instrument; therefore, numbers total more than the 11 studies represented, and percents total more than 100.

In seven studies, non-academic protocols or surveys developed by the authors of the studies were used. This was the most commonly-used type of instrument. Examples included demographic data sets (Kern et al., 2019), questionnaires with rating scales of the students’ testing experiences (Hahn et al., 2019), surveys of educators’ familiarity with accommodations policies (McCormack, 2019) and teachers’ perceptions and knowledge of accommodations (Carey et al., 2019), and observation protocols for measuring reading speed and accuracy (Giusto & Ehri, 2019).

Surveys or academic tests developed by researchers, or other education professionals, using sources outside of the study were employed in three studies. An example of a survey in the studies in 2019 was adapted by the researchers (Castro et al., 2019) from the Test Observation Form (TOF; McConaughy & Achenbach, 2004). An example of an academic test that was created is one by Robinson and colleagues (2019); they adapted selected grade level reading passages from AIMSweb, an assessment system (Shinn & Shinn, 2002).

State or district criterion-referenced assessments were employed in four studies. State tests were from New York (McCormack, 2019) and from three unidentified states (Buzick, 2019; Quesen & Lane, 2019), and the district test was from Portland, Oregon (Wise et al., 2019). Three norm-referenced academic achievement measures were used. For pre-screening or independent checking of performance, Giusto and Ehri (2019) used the Woodcock Reading Mastery Test-Revised (WRMT-R; Woodcock, 1987) and Kern and colleagues (2019) used the Woodcock Johnson Tests of Achievement, Third Edition (WJ-III; Woodcock et al., 2001). In addition, Giusto and Ehri employed the Gates MacGinitie Reading Comprehension Test, Fourth Edition, Form S, Grade 3 (GMRT; MacGinitie et al., 2002) for measuring a dependent variable. Two norm-referenced cognitive ability measures were applied for screening and identification purposes, both by Giusto and Ehri: the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4; Dunn & Dunn, 2007) and the Clinical Evaluation of Language Fundamentals, Fourth Edition

(CELF-4; Semel et al., 2003). In addition, Castro and colleagues (2019) used the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV; Wechsler, 2003) for measuring a dependent variable. A non-state criterion-referenced academic achievement measure on science, technology, engineering, and math (STEM) content was co-developed with an educator for students with visual impairments (Hahn et al., 2019). Over 60 percent of all studies ($n=7$) used instrumentation of more than one kind. A complete listing of the instruments used in each of the studies is provided in Table C-1 in Appendix C, including the related studies or other sources for these instruments, when available.

Content Area Assessed

Ten studies published during 2019 focused on accommodations used in specific academic content areas. As shown in Table 5, mathematics was the most commonly studied content area. Table 5 was constructed using data from the 2017 report (Rogers et al., 2020) and the 2018 report (Rogers et al., 2021). In all three years, reading and mathematics were the most common content areas for accommodations research, yet have varied in terms of which of the two was the most common content area in any particular year.

Cumulatively, science has tended to be the third most frequent content area, with five total studies across the three years of research reviews. In 2019, two-fifths of studies examined accommodations impact data for more than one content area. The inclusion of multiple content area analyses varied in frequency across the three years, from two (22%) in 2018 to four (40%) in 2019. The remaining academic content areas—writing, “other language arts,” and cognitive skills—comprised similar proportions of the research across the three years of studies. (See Appendix C, Table C-2, for additional details about the content areas.)

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

Content Area Assessed	2017 ^a	2018 ^a	2019 ^b
Mathematics	5 (45%)	7 (78%)	5 (50%)
Reading	6 (55%)	1 (11%)	4 (40%)
Writing	1 (9%)	0 (0%)	1 (10%)
Science	3 (27%)	1 (11%)	1 (10%)
Other language arts ^c	2 (18%)	1 (11%)	2 (20%)
Cognitive skills	0 (0%)	0 (0%)	1 (10%)
Multiple content ^d	4 (36%)	2 (22%)	4 (40%)
Total (of Relevant Studies)^e	11	9	10

^a Studies in 2017 and 2018 included studies that addressed more than one content area (i.e., two content areas, three content areas).

^b Studies in 2019 included studies that addressed more than one content area (i.e., two content areas).

^c Detailed descriptions of what constituted “Other Language Arts” for the 2019 studies can be found in Appendix C, Table C-2.

^d Because some studies investigated effects in more than one content area, the percents total more than 100.

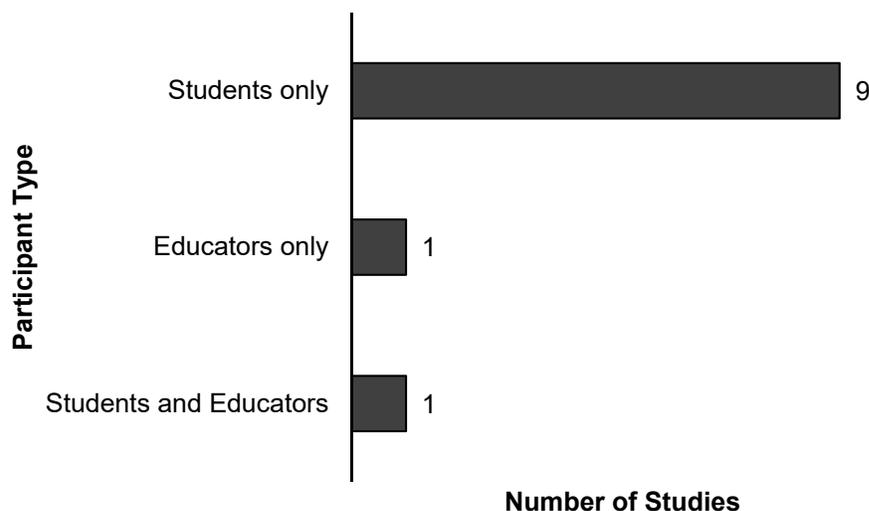
^e These totals were less than all studies analyzed from these years; in 2019, one study (Carey et al., 2019) did not address a specific content area.

Research Participants

As shown in Figure 3 and Appendix D, the studies in this review of accommodations research included students, educators, or both students and educators as participants. A majority of the studies included students only ($n=9$, 82%) and just one study included educators only ($n=1$, 9%). One of the studies in this review included both students and educators; none included parents.

Five studies (Giusto & Ehri, 2019; Hahn et al., 2019; Kern et al., 2019; Noakes et al., 2019; Robinson et al., 2019) specified the disabilities of the students included in their participant samples (see Appendix D). The nature of the disabilities was not specified in four of the studies (Buzick, 2019; Castro et al., 2019; McCormack, 2019; Quesen & Lane, 2019); in one study (Wise et al., 2019) only students without disabilities participated. Carey (2019) and McCormack both surveyed educators as a component of their studies that, in part, reported on the implementation and use of accommodations.

Figure 3. Types of Research Participants for K–12 Studies in 2019



The size and composition of the K–12 student participant groups in the 10 research studies that included students are shown in Table 6. For additional details about study participants, see Appendix D. In this set of reviewed studies, student participant groups varied from just three students (Noakes et al., 2019) to a large sample of 65,000 students (Buzick, 2019). The most common student group size was between 50 and 100 students ($n=3$), with the majority of studies ($n=6$) having fewer than 500 students in their samples. Table 6 shows that the majority of studies involving student participants ($n=7$) included between 75% and 100% students with disabilities. Only two studies included student participant groups that were less than 50% students with disabilities (Giusto & Ehri, 2019; Wise et al., 2019). Giusto and Ehri compared students with and without learning disabilities; however, only 34% of the students were students with disabilities.

None of the studies compared outcomes for equally-sized groups of students with and without disabilities. A single study, Wise et al., did not include any students identified with disabilities.

Table 6. Participant Sample Sizes and Ratio of K–12 Students with Disabilities in 2019

Number of Research Participants by Study	Number of Studies by Proportion of Sample Comprising Students with Disabilities				
	0–24%	25–49%	50–74%	75–100%	Total
1–9	0	0	0	1	1
10–49	0	0	0	1	1
50–99	0	1	0	2	3
100–499	0	0	0	1	1
500–999	0	0	0	0	0
1,000–9,999	0	0	0	1	1
10,000–19,999	1	0	0	1	2
20,000–100,000	0	0	1	0	1
Total	1	1	0	7	

School Level

This review of accommodations research identified 10 studies that included students at the elementary, middle, or high school levels (Table 7; see Appendix D for students’ specific grade levels when available). A majority of studies with student participants included multiple school levels ($n=6$). Two studies included students in just one grade level (Giusto & Ehri, 2019, grade 3; Quesen & Lane, 2019, grade 8) and a single study included participants in all grade levels (Wise et al., 2019, K–12).

The studies trended toward including participants in the earlier grades, with eight studies at the elementary level (73%) and six at the middle school level (55%), including the four studies composed of both elementary and middle school students. The four studies with student participants at the high school level (36%) also all incorporated students at the earlier grade levels (elementary, middle, or both). There were no studies that focused solely on high school students. Ten of the 11 studies had student participants, and one study had only educators as participants.

Table 7. School Level of Research Participants for K–12 Studies in 2019

School Level of All Participants	Number of Studies	Percent of Studies
Elementary school (K–5)	8	73%
Middle school (6–8)	6	55%
High school (9–12)	4	36%
Not applicable	1	9%

Note. Five studies (45%) had participants in more than one schooling level; therefore, the numbers total more than the 11 studies represented, and percents total more than 100.

Disability Categories

In 2019, the K–12 accommodations studies addressed a number of disability categories (see Appendix D for details). As Table 8 shows, the studies included students from several disability categories, without one category having been emphasized significantly more than the others. The largest proportion of the 11 studies focused on student participants with learning disabilities ($n=3$, 27%). Two studies (18%) did not include students with disabilities, while two studies did not specify the disability categories represented by the participating students. Additionally, the one study identified as “not applicable” did not include student participants.

In the only study that included students with visual impairments, Hahn and colleagues (2019) compared the outcomes for students in grades 5 through 8 when using touch screen devices and embossed paper as accommodations. Similarly, Noakes and colleagues (2019) were the only researchers to include students with traumatic brain injury (TBI); their focus was on comparing outcomes for students with TBI when using speech-to-text assistive technology on assessments. The students who participated in Castro and colleagues’ study (2019) were in pre-kindergarten through grade 5 and had been referred for consideration of special education services at the time of the study which compared two intelligence testing formats; Table 8 identifies these study participants as “other” because their status as having been referred for special education services was integral to the study.

Appendix D shows that two of the studies in this review included student participants from more than one disability category. Kern and colleagues (2019) investigated types of accommodations provided to high school students who exhibited emotional and behavioral problems and were identified for special education services. However, not all of the students were identified with an emotional behavioral disability; the sample included students identified with behavioral disabilities, learning disabilities, physical disabilities and speech/language impairments. The other study with students in more than one disability category is by Robinson and colleagues (2019). All of the student participants were elementary students identified with specific learning disabilities in reading. However, some of the students were also identified with attention

Table 8. Disabilities Reported for Research Participants for K–12 Studies in 2019

Disabilities of Research Participants	Number of Studies	Percent of Studies
Learning disabilities	3	27%
Speech/language impairments	2	18%
Attention problem	1	9%
Blindness/visual impairment	1	9%
Emotional behavioral disability	1	9%
Physical disability	1	9%
Traumatic brain injury	1	9%
Other ^a	1	9%
No disability	2	18%
Not specified	3	27%
Not applicable	1	9%

Note. Several studies had participants who fell into various disability categories; therefore, the numbers in this figure total more than the 11 studies represented, and percents total more than 100.

^a Other: at risk for developing disabilities.

problems and speech/language impairments. These studies referencing more than one disability category are included more than once in the data for Table 8 and account for the numbers of studies exceeding the 10 studies that included student participants and 11 total studies reviewed.

Types of Accommodations

The research studies published in 2019 are summarized in Table 9 according to the categories of accommodations being studied. Presentation was the most commonly investigated accommodation category ($n=8$). Response accommodations were examined in over one-third of the studies, and two studies each addressed accommodations in the remaining categories: equipment/materials, timing/scheduling, and setting.

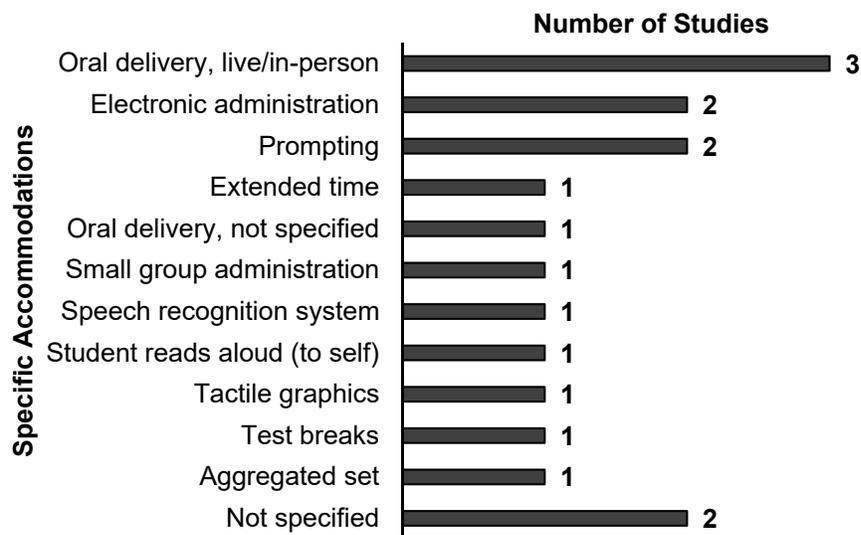
Table 9. Accommodation Categories for K–12 Studies in 2019

Accommodations Category	Number of Studies ^a
Presentation	8
Equipment/Materials	2
Response	4
Timing/Scheduling	2
Setting	2

^a Five studies investigated accommodations from more than one category; therefore, the numbers in this table total more than the 11 studies represented.

Figure 4 shows the specific accommodations investigated in the studies published in 2019. The most frequently studied presentation accommodation was oral delivery, in four studies (Giusto & Ehri, 2019; Kern et al., 2019; McCormack, 2019; Quesen & Lane, 2019). In previous NCEO accommodations research reports, we combined the three ways that the oral delivery accommodation was provided: (a) by a test administrator live and in-person, (b) with a recorded human voice, and (c) as simulated speech via text-to-speech devices or software. For additional information, see Table E-2 in Appendix E, which provides details about each of these oral delivery methods separately. Oral delivery presented live and in-person was investigated in two studies (McCormack; Quesen & Lane). Oral delivery presented live and in-person alone was combined, or bundled, with pacing by test administrators in one study (Giusto & Ehri), comparing this bundle to the pacing-only condition to discern the potential impact of in-person oral delivery alone. In another study describing several accommodations (Kern et al.), the manner in which oral delivery was provided was not specified.

Figure 4. Specific Accommodations for K–12 Studies in 2019



Note. Two studies each examined the separate impacts of more than one accommodation; therefore, the total exceeds the 11 studies represented.

Another presentation accommodation, prompting, was investigated by two studies (Giusto & Ehri, 2019; Wise et al., 2019). Giusto and Ehri (2019) made use of a comparison condition in which students were provided only pacing—that is, test proctors guiding students through the segments of the test but without reading items aloud. Performance while using this pacing-only accommodation was compared to using both pacing and partial oral delivery of the test, and to not using accommodations at all. Wise and colleagues (2019) examined the effect of test proctors checking in with students individually when the sensitive computer software detected student test-taking behavior associated with disengaged item responding such as rapid guessing.

Finally, the presentation and equipment/materials accommodation of electronic assessment administration was examined by two studies (Castro et al., 2019; Hahn et al., 2019). Castro and colleagues compared the performance of students with disabilities on a cognitive skills assessment presented on an electronic tablet to their performance on the assessment presented on paper. Hahn and colleagues crafted a similar performance comparison of two versions of the tactile graphics accommodation, providing graphics on paper for students with visual impairments including blindness and also on a multisensory electronic tablet that used haptic features (vibration).

Response accommodations were investigated in four studies: two were the previously-mentioned electronic administration studies (Castro et al., 2019; Hahn et al., 2019), one studied calculator use and dictated response (Kern et al., 2019), and one studied a speech recognition system (Noakes et al., 2019).

The timing/scheduling accommodation of test breaks was part of the accommodations package in one study (Kern et al., 2019), and was examined separately in one study (Quesen & Lane, 2019). The setting accommodation of small group was part of the accommodations package in one study (Kern et al., 2019), and was examined separately in one study (Quesen & Lane, 2019).

Five studies included accommodations from more than one category. Of those, one study included accommodations from each of four accommodations types, two studies included accommodations from each of three accommodations types, and two studies included accommodations from two accommodations types. A complete listing of accommodations examined in each study is provided in Appendix E Table E-1, and by accommodation type in Tables E-2 through E-6.

Research Findings

The findings of the studies on accommodations published in 2019 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 oral delivery. One study examined impacts of aggregated sets of accommodations sometimes called “bundles.” We also present findings on the impact of unique accommodations—those examined in only one study—such as reading aloud to oneself or assistive technology (see Appendix F). This report includes perceptions of accommodations, including those of student test-takers and of educators. We summarize the findings on the effects of accommodations, and describe implementation conditions as well as patterns of use of various accommodations. This report also presents findings by academic content areas: math, reading, and other language arts. Details on individual studies are available in Appendix F.

Impact of Accommodations

Research published in 2019 that examined the effects of accommodations on assessment performance for K–12 students with disabilities totaled six studies (see Appendix F for details about each of these studies). We report here on the effects of two discrete accommodations: electronic administration and prompting. See Appendix F for further details on effects of each accommodation with only one associated study.

Two accommodations examined for their effects on assessment performance were investigated in two studies each: electronic administration (Castro et al., 2019; Hahn et al., 2019) and prompting (Giusto & Ehri, 2019; Wise et al., 2019). **Electronic administration** was provided in comparison with standard paper format (Castro et al., 2019) or in a comparison of a multisensory tablet to embossed paper graphic images (Hahn et al., 2019). **Prompting** was presented as a combination of pacing support with partial oral delivery compared to pacing only (Giusto & Ehri, 2019), as well as a combination of engagement monitoring and proctor notification to provide prompting compared to no monitoring and no notification for prompting (Wise et al., 2019).

Castro and colleagues (2019) examined the performance of elementary students on a cognitive skills assessment when it was administered either on a tablet or in a typical paper-and-pencil format. The students in the study all were identified with various disabilities or conditions associated with testing difficulties (e.g., depression, anxiety, behavioral disabilities, attention problems, language/thought problems). The performance of the two groups, on average, was similar regardless of the format of administration. Hahn and colleagues (2019) examined the performance of students with visual impairments in grades 5–12 on math/science items using either multisensory tablet that had a touchscreen with tactile haptic features (vibration) or embossed paper tactile graphic images. Results of paired sample mean scores showed no significant differences in performance, even though there were slightly more correct answers with the embossed paper presentation. The severity of the students' visual impairments, their experience with braille reading, and their previous experience with the tablet did not mediate the results.

Giusto and Ehri (2019) found that grade 3 students with learning disabilities who were provided both prompting support (in the form of pacing) and partial oral delivery scored significantly higher on reading comprehension questions than students with learning disabilities provided only pacing support and students provided no accommodations. Oral delivery was provided only for test instructions and item stems, not for reading passages. The students without learning disabilities did not score differently across the three testing conditions. Wise and colleagues (2019) found higher test performance in K–12 students when computer-based math and reading tests were monitored for test engagement and proctors were notified of disengagement (defined as three consecutive rapid guesses). Proctors who were notified were allowed to follow up with the student in the way deemed appropriate by the proctor. The positive effects of monitoring and

proctor notification were stronger for reading tests than for math tests. The authors also found that the correlations between beginning test performance and final test scores were higher for students in the engagement monitoring and proctor notification conditions. Proctors generally viewed the proctor notification feature as valuable.

We identified separate findings on the impact of two **unique accommodations**—that is, accommodations that were the focus of just one study. Effects of these two unique accommodations were examined in two studies:

- speech recognition system (Noakes et al., 2019)
- reading aloud to self (Robinson et al., 2019)

Noakes and colleagues (2019) studied the writing performance of three students with traumatic brain injury (grades 4, 8, and 9) when using speech-to-text assistive technology and when using only handwriting in response to prompts. Performance was measured in terms of total words written, words spelled correctly, and correct writing sequences. Results indicated that students scored higher in the speech-to-text condition, with non-overlapping scores for the two conditions. Large effect sizes were found for all students favoring the speech-to-text condition. Robinson and colleagues (2019) examined the effects of the student reading aloud to self in comparison to the student reading silently during an assessment of reading comprehension for students with reading-related disabilities. The 77 student participants from grades 2–5 were identified as having attention problems, learning disabilities, or speech/language impairment. Results indicated that students in grades 2 and 3 had better retell performance when they had read aloud to themselves than when they read silently. In contrast, students in grades 4 and 5 had similar performance for the two conditions. Findings for each of these unique accommodations are presented in Appendix F.

Use and Implementation of Accommodations

Four studies (Buzick, 2019; Carey et al., 2019; Kern et al., 2019; McCormack, 2019) had findings related to accommodations use and implementation issues. In two studies (Buzick, 2019; Kern et al., 2019), researchers described patterns of accommodations use, while two studies (Carey et al., 2019; McCormack, 2019) provided information about educators' accommodations implementation practices.

Buzick (2019) reported that within their sample most students with disabilities in grades 3 through 8 from two states were assigned accommodations in two consecutive years, yet approximately 25% were not. The study found that the second most frequent pattern of accommodations assignment was not receiving accommodations in the first year but receiving them in the second year. Two analytic methods yielded correlations between inconsistent use of accommodations

across academic years and different rates of learning growth in English language arts and math. Students who were assigned accommodations only in the second year averaged the largest academic performance improvement compared to students with other use patterns. Kern and colleagues (2019) analyzed data for secondary students with disabilities from five states and described patterns of accommodations use in the classroom setting and during state assessments. Classroom supports included test-related accommodations: 48% received extended test time, 48% received small group testing, and 41% received oral delivery. State/district assessment accommodations included extended time for 56% of students, small group for 90%, and oral delivery for 52%. The number and category of accommodations provided to students with disabilities tended not to be associated with race/ethnicity, gender, or grade level.

Carey and colleagues (2019) reported survey findings on assessment accommodations practices and knowledge, as well as beliefs or perceptions, for general and special education teachers at elementary, middle, and high school levels. Educators reported practices that were strongly associated with beliefs and knowledge. Special education teachers supported practices such as modeling and discussing accommodations with students to ensure that they benefited. General education teachers, especially at the high school level, tended not to do so, and indicated a lack of confidence in their knowledge and ability to appropriately provide accommodations. McCormack (2019) reported the results of a survey of 37 special education administrators from across New York State. Respondents indicated that the oral delivery accommodation was assigned to “level the playing field” rather than to provide access. The survey also revealed a lack of knowledge or use of the state’s accommodations guidelines or decision-making tool, and respondents reported limited training for teachers on accommodations in general, and specifically on the state’s decision-making tool.

Perceptions about Accommodations

Two studies (Carey et al, 2019; Hahn et al., 2019) provided findings on perceptions about accommodations. Carey and colleagues examined the beliefs and knowledge of 240 educators, both special education teachers and general education teachers, and the relationship of those to their accommodation practices. Hahn and colleagues explored the perceptions of 22 students with visual impairments in grades 5–12 about the use of a multisensory tablet that provided a touchscreen with haptic features (vibrations) and sound, compared to embossed paper tactile graphic images on math and science items.

Carey et al. (2019) found that teacher beliefs, knowledge, and practice responses were generally positive, with beliefs significantly higher than knowledge and practice, and also that beliefs and knowledge were significantly associated with practice. Further, there were differences in the beliefs and knowledge of special education teachers and general education teachers, with the former having the most positive beliefs about accommodations. Even though perceptions of ac-

accommodations were generally positive, all teachers (and especially general education teachers) indicated that they lacked confidence in their knowledge and ability to appropriately provide accommodations and in their ability to instruct students on accommodation use.

Hahn et al. (2019) sought the perceptions of students who participated in their study on the effects of tactile graphics presented through embossed paper or through a multisensory tablet. The responding students provided both positive and negative feedback regarding the multisensory tablet. Among the positive feedback was that the students enjoyed using the device; they thought the sounds and vibrations were engaging. Negative feedback included concerns about the use of space, line thickness, and repetitive sounds. Additional details about the perception findings of the Carey et al. (2019) and Hahn et al. studies are provided in Appendix F.

Accommodations by Academic Content Assessments

We analyzed research findings according to the academic content area included in each of the studies in 2019, as we have in previous reports (Rogers et al., 2012, 2014, 2016, 2019, 2020, 2021). The content areas, presented in terms of the number of studies (when $n > 1$) are: mathematics ($n=5$), reading ($n=4$), and other English language arts (ELA; $n=2$). For each content area, this report presents a summary of findings for the impact of accommodations on assessment performance, perceptions about accommodations, implementation matters and use patterns, and approaches to calculating item response patterns in state assessment data sets. This content-related discussion addresses the eight studies with the three content areas (mathematics, reading, other ELA); accommodations for each of these content areas were addressed in more than one study. One of these studies (Hahn et al., 2019) incorporated findings pertaining to both math and science. Two studies not discussed here investigated accommodations in either writing (Noakes et al., 2019) or cognitive skills (Castro et al., 2019). (See Appendix F for a more detailed explanation of the findings of each study.)

Mathematics. Five studies (Buzick, 2019; Hahn et al., 2019; Kern et al., 2019; Quesen & Lane, 2019; Wise et al., 2019) presented research findings on accommodations for mathematics assessments. Two studies presented findings on the effects of specific supports on math test performance, including tactile graphics on paper versus sound and haptic graphics on electronic tablet for students with visual impairments (Hahn et al., 2019) and prompting support for students without disabilities to re-engage in testing (Wise et al., 2019). In contrast, one study (Buzick, 2019) reported on degrees of academic growth in math (and English language arts) based on unspecified accommodations in general in only one or in two consecutive years. One study (Kern et al., 2019) reported on the types of accommodations provided on classroom and state/district assessments but did not directly examine performance effects. Finally, Quesen and Lane (2019) explored use of several accommodations—extended time; oral delivery live/in-person, small group, test breaks—when describing item-level responses on math assessment data sets.

Four of the five studies used extant data sets for the analyses, with only one study (Hahn et al., 2019) presenting primary data.

The use of supports versus no supports benefited the math performance of students without disabilities in one study (Wise et al., 2019), while in the other study (Hahn et al., 2019) one version of accommodations was more beneficial for students with disabilities than the other version. Wise and colleagues found that test engagement monitoring and proctor notification were beneficial and resulted in higher student engagement and higher performance, although the effects for math were less pronounced than for reading. Hahn and colleagues indicated that students were more accurate by 6% when using the embossed graphics on paper in comparison to when using the multisensory presentation of math graphics on a tablet with sound and haptics; however, this difference was not significant.

Two studies (Buzick, 2019; Kern et al., 2019) inspected extant data sets for patterns of accommodations provision. Buzick analyzed statewide data sets across two states for the consistency of providing unspecified accommodations to students in grades 4–8 in math (and ELA). Proportions of students with disabilities assigned accommodations for math in only one year ranged from 4–25% across grade levels; 10–32% were assigned no accommodations for math (presumably because they did not need them). Kern and colleagues examined the accommodation types provided to students in the classroom and on state/district assessments in math (and reading). During both classroom and state assessments, students with lower math performance—according to the Woodcock Johnson Tests of Achievement, Third Edition (WJ-III; Woodcock et al., 2001)—were more likely to be assigned response accommodations, including calculators. During state assessments, students with lower math (and reading; per WJ-III) scores had more likelihood of being assigned presentation accommodations.

Only one math-related study (Hahn et al., 2019) presented findings on perceptions about accommodations. This observation is quite different from a previous NCEO report describing findings of studies from 2018, in which nearly all seven math-related accommodations studies provided findings on accommodations' effects on performance and also yielded study participants' perceptions regarding the accommodations (Rogers et al., 2021). When asked about their experiences using a multisensory tablet with both sound and haptic features to complete math (and science) test items, students with visual impairments provided both positive and negative feedback. Participants expressed their enjoyment and self-reported higher engagement, yet also commented about technological complications with sensing line thickness (Hahn et al.).

Reading. Four studies (Giusto & Ehri, 2019; Kern et al., 2019; Robinson et al., 2019; Wise et al., 2019) provided research findings on accommodations for reading assessments. Three studies presented findings on the effects of specific accommodations on reading test performance, including prompting alone versus prompting and oral delivery for students with reading dis-

abilities (Giusto & Ehri, 2019), students with reading disabilities reading aloud to themselves (Robinson et al., 2019), and prompting students with no identified disabilities to stay on task when appearing not engaged (Wise et al., 2019). Giusto and Ehri found that students performed better in reading comprehension when provided by test proctors with both prompting support and partial oral delivery than either with prompting alone or without accommodations. Students using only prompting support did not score significantly differently than students not using accommodations. In contrast, students without disabilities did not score significantly differently across the three testing conditions. Robinson and colleagues reported that students in grades 2 and 3 had better retell performance when they had read aloud to themselves than when they read silently, yet students in grades 4 and 5 exhibited no performance differences when using or not using the accommodation. Wise and colleagues found that test engagement monitoring and proctor notification were beneficial and resulted in higher student engagement and higher performance, with more positive effects for reading than for math. One study (Kern et al.) analyzed the categories of accommodations provided during classroom and state/district assessments in reading (and math). During state assessments, students with lower reading performance—according to the Woodcock Johnson Tests of Achievement, Third Edition (WJ-III; Woodcock et al., 2001)—were more likely to receive presentation accommodations, including oral delivery.

Other Language Arts. Two studies (Buzick, 2019; McCormack, 2019) provided findings for other English language arts (ELA). Neither study reported conclusively on the impact of accommodations on test performance. Instead, Buzick illuminated accommodations use patterns by elementary and middle school students in a dataset of two statewide assessments of English language arts (and math). Most students were assigned accommodations in both years. The second most frequent pattern was not receiving accommodations in the first year but receiving them in the second year. Students who were assigned accommodations only in the second year averaged the largest ELA performance improvement. McCormack described teacher survey results on their accommodations perceptions, knowledge, and practices, along with an analysis of an extant dataset of assessment performance by grade 4 students with disabilities on an English language arts assessment based on the NYS P–12 Common Core Learning Standards (Questar Assessment, 2016). Teacher surveys revealed several issues, including gaps in knowledge about or application of the state’s accommodations guidelines or decision-making tool. ELA proficiency levels for grade 4 students with various disabilities were significantly different among regions of New York, and the use of the live/in-person oral delivery accommodation ranged widely across school districts; however, post hoc tests did not indicate that oral delivery was associated with test score differences. (See Appendix F for more detailed explanations of the findings of each study.)

Discussion

This is the third consecutive NCEO report that has provided a snapshot of accommodations research literature involving K–12 students published within a single year; this report addresses studies from 2019. This is a narrower focus than previous NCEO accommodations research reports (Cormier et al., 2010; Johnstone et al., 2006; Thompson et al., 2002; Rogers et al., 2012; Rogers et al., 2014; Rogers et al., 2016; Rogers et al., 2019; Zenisky et al., 2007), which have each encompassed more than one year of research literature, and incorporated studies from elementary, secondary, and postsecondary education settings. This report addresses the types of accommodations that were studied, the purposes of the research, the research type, data sources, characteristics of the independent and dependent variables under study, and the comparability of findings between studies in similar domains, including for specific accommodations and their performance effects, and examined findings by academic content area.

Mathematics and reading were the content areas most frequently examined in the studies included in this analysis—together comprising most of the studies that used content assessments. Science comprised exactly one-tenth of the studies with content assessments. These proportions are generally similar to those noted in previous NCEO accommodations research reports (Cormier et al., 2010; Johnstone et al., 2006; Thompson et al., 2002; Rogers et al., 2012; Rogers et al., 2014; Rogers et al., 2016; Rogers et al., 2019; Rogers et al., 2020; Rogers et al., 2021; Zenisky et al., 2007). An exception is that math alone comprised a larger proportion of studies, while reading, writing, or English language arts were together addressed in an equivalent proportion.

Students were the participant group in nearly all of the studies from 2019. Students with learning disabilities were participants in the largest proportion, about one-third, of the studies reported, which is typical for this series of accommodations research reports. Other disability categories receiving attention by some studies included attention-related disabilities, physical and mobility disabilities, speech/language impairment, traumatic brain injuries, visual impairments including blindness, intellectual disabilities, and multiple disabilities; each of these categories were part of one or two studies each. Three studies—more studies than is typical in a single study year—reported information on the population of students with disabilities as a whole, without identifying data for students from specific disability categories.

Research into the various aspects of assessment accommodations has been a continually expanding inquiry, generating attention in both breadth and depth. The number of studies we have identified has increased across the span of NCEO's reports in this area; for instance, in 2011–2012, there were 49 identified studies, in 2013–2014, there were 53 studies, and in 2015–2016, there were 58 studies. A substantial proportion of the studies examined in those previous reports highlighted research on the academic accommodations provided at institutions of higher education, and several have been from K–12 and postsecondary contexts outside the

United States. Keeping in mind that the past two NCEO reports in this series (Rogers et al., 2020; Rogers et al., 2021) each addressed only one year (2017 and 2018, respectively) and were narrowed to the context of the U.S. K–12 school system, the 25 studies described in these two years were consistent with a continually increasing trend for this topic in research literature. Further, the current report also addressed one year, 2019, of accommodations research within the U.S. K–12 context, and 11 studies were examined.

Researchers have continued to explore a wide range of topics related to assessment accessibility features and accommodations, including the comparison of the effects of differing versions of accommodations, such as providing tactile graphics either on paper or electronically with haptics and sound. Researchers simultaneously addressed multiple study purposes and examined various types of data within some studies. Along with investigating effects of accommodations during content tests, researchers inquired about perceptions of students with disabilities and educators about accommodations, as well as students' use of, and educators' implementation practices pertaining to, accommodations. These investigations have yielded findings that can inform implementation of emerging accessibility tools for students with disabilities, and can provide new knowledge for considering ways to address students' needs. Further, issue-driven inquiries can offer additional insights and perspectives about the provision of assessment accommodations to students with disabilities who need them.

Report References

(References in the report to studies that were part of the 2019 accommodations research analysis are not included in this list. They are in the separate list titled 2019 K–12 Accommodation References.)

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). *The Peabody picture vocabulary test* (4th ed.; PPVT™-4). Pearson. <https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Academic-Learning/Brief/Peabody-Picture-Vocabulary-Test-%7C-Fourth-Edition/p/100000501.html>

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.

Leslie, L., & Caldwell, J. A. (2011). *Qualitative reading inventory* (5th ed.). Pearson Education.

MacGinitie, W., MacGinitie, R., Maria, K., & Dreyer, L. (2002). *The Gates MacGinitie reading comprehension test* (4th ed.). Riverside.

March, J. (1998). *Manual for the multidimensional anxiety scale for children* (MASC). Multi-Health Systems. <https://doi.org/10.1037/t05050-000>

McConaughy, S. H., & Achenbach, T. M. (2004). *Manual for the test observation form for ages 2–18*. Research Center for Children, Youth & Families.

Powell-Smith, K. A., & Shinn, M. R. (2004). *Administration and scoring of written expression curriculum-based measurement (WE-CBM) for use in general outcome measurement*. Edformation.

Questar Assessment. (2016). *New York state testing program 2016: English language arts and mathematics grades 3–8* (Technical Report). New York State Education Department. <https://eric.ed.gov/?q=%22New+York+state+testing+program+2016%3a+English+language+arts+and+mathematics+grades+3%e2%80%938%22&id=ED591460>

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior assessment system for children* (2nd ed.). American Guidance Service. <https://doi.org/10.1037/t15079-000>

Reynolds, W. M. (2002). *Reynolds adolescent depression scale: Professional manual* (2nd ed.). Psychological Assessment Resources.

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., 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.

Semel, E., Wiig, W. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals* (4th ed.). Pearson.

Shinn, M. M., & Shinn, M. R. (2002). *AIMSweb training workbook: Administration and scoring of reading curriculum-based measurement (R-CBM) for use in general outcome measurement*. Edformation.

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. (2003). *Wechsler intelligence scale for children* (4th ed.; WISC-IV). The Psychological Corporation. https://doi.org/10.1007/springreference_70197

Wise, S. L., & DeMars, C. E. (2006). An application of item response time: The effort-moderated IRT model. *Journal of Educational Measurement*, 43, 19–38. <https://doi.org/10.1111/j.1745-3984.2006.00002.x>

Woodcock, R. W. (1987). *Woodcock reading mastery tests-revised*. American Guidance Service. <https://doi.org/10.1037/t15179-000>

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Riverside Publishing.

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.

2019 K–12 Accommodation References (N=11)

Buzick, H. M. (2019). Testing accommodations and the measurement of student academic growth. *Educational Assessment*, 24(1), 57–72. <https://doi.org/10.1080/10627197.2018.1545571>

Carey, L. B., Stephan, C., & Pritchard, A. E. (2019). Preparing students for competent use of academic testing accommodations: Teachers’ belief, knowledge, and practice. *Learning Disabilities: A Multidisciplinary Journal*, 24(1), 33–42. <https://doi.org/10.18666/LDMJ-2019-V24-I1-9021>

Castro, C. J., Viezel, K., Dumont, R., & Guiney, M. (2019). Exploration of children’s test behavior during iPad-administered intelligence testing. *Journal of Psychoeducational Assessment*, 37(1), 3–13. <https://doi.org/10.1177/0734282917729304>

Giusto, M., & Ehri, L. C. (2019). Effectiveness of a partial read-aloud test accommodation to assess reading comprehension in students with a reading disability. *Journal of Learning Disabilities*, 52(3), 259–270. <https://doi.org/10.1177/0022219418789377>

Hahn, M. E., Mueller, C. M., & Gorlewicz, J. L. (2019). The comprehension of STEM graphics via a multisensory tablet electronic device by students with visual impairments. *Journal of Visual Impairment & Blindness*, 113(5), 404–418. <https://doi.org/10.1177/0145482X19876463>

Kern, L., Hetrick, A. A., Custer, B. A., & Commisso, C. E. (2019). An evaluation of IEP accommodations for secondary students with emotional and behavioral problems. *Journal of Emotional and Behavioral Disorders*, 27(3), 178–192. <https://doi.org/10.1177/1063426618763108>

McCormack, A. (2019). *An investigation into the use of the “tests read” accommodation for the New York State 4th grade ELA assessment, the impact on reported scores, and the reported knowledge and beliefs of administrators regarding the assignment and purpose of accommoda-*

tions (Publication No. 13899837) [Doctoral dissertation, St. John's University (NY)]. ProQuest Dissertations and Theses Global.

Noakes, M. A., Schmitt, A. J., McCallum, E., & Schutte, K. (2019). Speech-to-text assistive technology for the written expression of students with traumatic brain injuries: A single case experimental study. *School Psychology, 34*(6), 656–664. <https://doi.org/10.1037/spq0000316>

Quesen, S., & Lane, S. (2019). Differential item functioning for accommodated students with disabilities: Effect of differences in proficiency distributions. *Applied Measurement in Education, 32*(4), 337–349. <https://doi.org/10.1080/08957347.2019.1660347>

Robinson, M. F., Meisinger, E. B., & Joyner, R. E. (2019). The influence of oral versus silent reading on reading comprehension in students with reading disabilities. *Learning Disability Quarterly, 42*(2), 105–116. <https://doi.org/10.1177/0731948718806665>

Wise, S. L., Kuhfeld, M. R., & Soland, J. (2019). The effects of effort monitoring with proctor notification on test-taking engagement, test performance, and validity. *Applied Measurement in Education, 32*(2), 183–192. <https://doi.org/10.1080/08957347.2019.1577248>

Appendix A

Research Characteristics for K–12 Studies in 2019

Authors	Publication Type	Research Type	Research Design	Data Collection Source	Collection Instrument
Buzick	Journal	Quantitative	Longitudinal	Secondary	Test
Carey et al.	Journal	Quantitative	Descriptive Quantitative	Primary	Survey
Castro et al.	Journal	Quantitative	Quasi-experimental	Primary	Survey, Test
Giusto & Ehri	Journal	Quantitative	Experimental	Primary	Observations, Test
Hahn et al.	Journal	Quantitative	Experimental	Primary	Survey, Test
Kern et al.	Journal	Quantitative	Longitudinal	Secondary	Survey, Test, Other
McCormack	Dissertation	Mixed	Descriptive Quantitative	Primary	Survey, Test
Noakes et al.	Journal	Quantitative	Experimental	Primary	Observations, Test
Quesen & Lane	Journal	Quantitative	Correlation/ Prediction	Secondary	Test
Robinson et al.	Journal	Quantitative	Quasi-experimental	Primary	Test
Wise et al.	Journal	Quantitative	Quasi-experimental	Secondary	Observations, Test, Other ^a

^a For further “Other” collection instrument information, see Appendix C, Table C-1.

Appendix B

Research Purposes for K–12 Studies in 2019

Authors	Stated Research Purpose	Effects (SwD)	Effects (non)	Effects (both)	Implementation/Use	Perceptions	Test Items	Validity	Issues	Research
Buzick	Describe accommodations use patterns and estimate inconsistencies' relation with student growth.				P					
Carey et al.	Inquire about teacher perceptions, beliefs, and knowledge, and their potential influence on the assignment of accommodations.				X	P				
Castro et al.	Compare the effects of administering intelligence tests via different formats—specifically standard paper format versus tablet—on the performance and behavior during testing of students referred for special education screening.	P								
Giusto & Ehri	Investigate test validity of a reading comprehension assessment when a partial read-aloud accommodation is used for poor decoders; compare the effects of a partial read-aloud accommodation on the reading comprehension scores of students identified and not identified as poor decoders.	X						P		
Hahn et al.	Compare the effects of the use of multimodal touch screen devices on the accessibility of science, technology, engineering, and mathematics graphics for students who are visually impaired; study how students with visual impairments perceive graphics presented via multimodal touch screen devices versus embossed paper format.	P				X				
Kern et al.	Report on the types of accommodations received by high school students with emotional and behavioral problems; discuss additional factors associated with accommodations use patterns.				P				X	

Authors	Stated Research Purpose	Effects (SwD)	Effects (non)	Effects (both)	Implementation/Use	Perceptions	Test Items	Validity	Issues	Research
McCormack	Report on regional implementation of the “tests read” accommodation by students with disabilities across regions of New York State; discuss issues related to special education leaders’ implementation and use of “tests read” accommodation; summarize the research regarding the “tests read” accommodation.				P				X	X
Noakes et al.	Investigate the effects of speech-to-text assistive technology (AT) on assessment scores for students with traumatic brain injury.	P								
Quesen & Lane	Examine extant item-level score data for differential item functioning (DIF), using different models of analysis, related to ability distributions.						P			
Robinson et al.	Compare the effects of students reading passages orally or silently on reading comprehension scores for elementary students with specific learning disabilities.	P								
Wise et al.	Investigate the effects of effort monitoring with proctor notification on student engagement and performance.		P							

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]
Implementation/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
Issues	Discuss issues related to test accommodations
Research	Summarize research on test accommodations
P	Primary purpose
X	Other purpose

Appendix C

Instrument Characteristics for K–12 Studies in 2019

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

Authors	Instrument Types and Description/s	Number of Types
Buzick	State Test: Extant score data from 2005–2009 administrations of two unidentified states' grades 3–8 ELA and mathematics tests	1
Carey et al.	Author (survey): The Knowledge, Beliefs, and Practices around Student Testing Accommodations Survey, developed with feedback from special education and psychometrics experts; 17 items: 5 knowledge scale items, 4 items on beliefs, and 8 on practices; used Likert scale from 1 (strongly disagree) to 5 (strongly agree); also collected demographics, training, and work experience details	1
Castro et al.	Researcher (survey): Test Observation Form (TOF; McConaughy & Achenbach, 2004) syndrome scales and broad scales for documentation of behaviors during test participation Norm-ref Ability: Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV; Wechsler, 2003) for full-scale intelligence quotient (IQ)	2
Giusto & Ehri	Author (observations): Observed test response time, recorded in number of words read silently and words read aloud Norm-ref Ach: For screening/identification: Word Identification Subtest and Word Attack Subtest on Woodcock Reading Mastery Test-Revised (WRMT-R; Woodcock, 1987). For primary dependent variable: Reading comprehension on the Gates MacGinitie Reading Comprehension Test, Fourth Edition, Form S, Grade 3 (GMRT; MacGinitie, MacGinitie, Maria, & Dreyer, 2002) Norm-ref Ability: For screening/identification: Norm-ref Ability: Receptive vocabulary on the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4; Dunn & Dunn, 2007). Understanding Spoken Paragraphs Subtest and Understanding Concepts and Spoken Directions Subtest on the Clinical Evaluation of Language Fundamentals, Fourth Edition (CELF-4; Semel et al., 2003)	3
Hahn et al.	Author (surveys): Self-report questionnaire on demographics, history of related disabilities, degree of impairment, primary learning modality, and preferred operating system; exit questionnaire seeking feedback on touch screen graphics, along with scaled items seeking scaled responses on problems with and impact of delayed access or inaccessibility to content, and motivational factors Crit-ref Ach: Set of questions measuring STEM (science, technology, engineering, math) content comprehension after using different accommodations, co-developed with an educator for students with visual impairments	2

Authors	Instrument Types and Description/s	Number of Types
Kern et al.	<p>Author (survey): Parents reported student participants’ demographic data including gender, race, grade level, primary disability category, and state of residence</p> <p>Norm-ref Ach: Broad reading and broad mathematics on the Woodcock Johnson Tests of Achievement, Third Edition (WJ-III; Woodcock, McGrew, & Mather, 2001)</p> <p>Other: For screening/identification: Internalizing or externalizing composites of the Behavior Assessment System for Children, Second Edition—Teacher or Parent Version (BASC-2; Reynolds & Kamphaus, 2004); Multidimensional Anxiety Scale for Children (MASC; March, 1998); Reynolds Adolescent Depression Scale, Second Edition (RADS-2; Reynolds, 2002); other data artifacts (for data triangulation), including IEP documents</p>	3
McCormack	<p>Author (survey): Educator survey’s purpose was to check special education leaders’ knowledge on the “test read” testing accommodation, including on district accommodations policies; 4 of the 5 questions were open-ended, regarding professional development and assignment of testing accommodations, along with one job role question; survey developed by author and revised based on feedback from five experts in the field of testing or research. Data on the number of students taking exams and the number with accommodations per district were collected separately from the state.</p> <p>State Test: Extant test score data from the 2017 and 2018 New York State 4th Grade ELA Exam (Questar Assessment, 2016), a criterion referenced multiple-choice and constructed-response test based on the NYS P–12 Common Core Learning Standards</p>	2
Noakes et al.	<p>Author (observations): Qualitative descriptions of participants’ handwriting, in terms of size and legibility, according to authors’ observations in comparing their handwriting with their student peers</p> <p>Researcher (test): The AIMSweb assessment system (Powell-Smith & Shinn, 2004) supplied 10 different story starter writing prompts, measuring written expression by total words written, words spelled correctly, and correct writing sequences</p>	2
Quesen & Lane	<p>State Test: Extant score data from an unidentified state’s grade 8 mathematics assessment</p>	1
Robinson et al.	<p>Researcher (tests): AIMSweb reading passages for grades 2, 3, 4, and 5 (Shinn & Shinn, 2002), shortened in length to attend to reader fatigue; participants’ words read (aloud and silently) were documented. For reading comprehension, the authors used a Reading Comprehension One Minute Retell, modeled on the Qualitative Reading Inventory, 5th edition (QRI, Leslie & Caldwell, 2011); that is, authors tallied the passages for number of idea units conveyed; participants provided a one-minute retell of the content they each completed.</p>	1

Authors	Instrument Types and Description/s	Number of Types
Wise et al.	<p>Author (observation): Observed test item response times</p> <p>District (large-scale) Test: District-wide (Portland, Oregon) computerized adaptive testing of interim achievement in math and reading; Fall 2016 administration and Fall 2017 administration. Note: The test events were not matched, but rather were checked with proxy variables and researchers deemed them reasonably equivalent.</p> <p>Other: Estimate of rapid-guessing threshold by test item, drawing from author’s independent research (Wise & DeMars, 2006) and seeking to serve as an indicator of test effort monitoring and engagement, addressing motivation</p>	3

KEY for Table C-1

Instrument Types	Type Abbreviations	Number of Studies
Non-Academic Protocols or Surveys Developed by Study Author/s	Author Survey / Interview / Protocol	7
State Criterion-referenced Assessment	State Test	4
Surveys or Academic Tests Developed by Professionals or Other Researchers through Work Outside of Current Study	Researcher Test	3
Norm-referenced Academic Achievement Measures	Norm-ref Ach	2
Norm-referenced Cognitive Ability Measures	Norm-ref Ability	2
Criterion-referenced Academic Achievement Measures	Crit-ref Ach	1
Other	Other	2

Table C-2. Content Areas Assessed

Authors	Math	Reading	Writing	Other LA	Science	Social Studies	Cognitive Skills	N
Buzick	•			• ^a				2
Castro et al.							•	1
Giusto & Ehri		•						1
Hahn et al.	•				•			2
Kern et al.	•	•						2
McCormack				• ^b				1
Noakes et al.			•					1
Quesen & Lane	•							1
Robinson et al.		•						1
Wise et al.	•	•						2
TOTAL	5	4	1	2	1	0	1	

^a In this study, other LA = identified by two unnamed states as English language arts.

^b In this study, other LA = identified by New York State Education Department and Questar Assessment (2016) as English language arts.

Appendix D

Participant and Sample Characteristics for K–12 Studies in 2019

Authors	Unit of Analysis	Sample Size	Percent of Sample with Disabilities	Grade / Education Level	Disability Categories Included in Sample
Buzick	Students	65,000	100%	Grades 4, 5, 6, 7, 8	Not specified
Carey et al.	Educators	240	N/A	N/A	N/A
Castro et al.	Students	93	100%	Grades preK, K, 1, 2, 3, 4, 5	Other
Giusto & Ehri	Students	82	34%	Grade 3	LD; None
Hahn et al.	Students	22	100%	Grades 5, 6, 7, 8, 9, 10, 11, 12	VI
Kern et al.	Students	222	100%	Grades 8, 9, 10, 11	EBD, LD, PD, S/L
McCormack	Educators, Students	3,281	99%	Grade 4	Not specified
Noakes et al.	Students	3	100%	Grades 4, 8, 9	TBI
Quesen & Lane	Students	18,795	68%	Grade 8	Not specified
Robinson et al.	Students	77	100%	Grades 2, 3, 4, 5	AP, LD, S/L
Wise et al.	Students	10,353	0%	Grades K–12	None

KEY for Appendix D

AP=Attention Problem

EBD=Emotional/Behavioral Disability

LD=Learning Disability

PD=Physical Disability

S/L=Speech/Language Impairment

TBI=Traumatic Brain Injury

VI=Visual Impairment/Blindness

Other=At Risk for Developing Disabilities

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 2019

Table E-1. All Accommodations by Study

Author/s	Accommodation/s
Buzick	Not specified
Carey et al.	Not specified
Castro et al.	Electronic administration
Giusto & Ehri	Prompting; oral delivery, live/in-person
Hahn et al.	Tactile graphics, on paper along with braille; multisensory graphics—tactile and haptics, presented on electronic tablet
Kern et al.	Aggregated set: Calculator; cueing; dictated response; extended time; format; individual; large print/magnification; mark answer in test booklet; oral delivery, not specified; physical supports; reinforcement; simplified language; small group; specialized setting; speech recognition system; technological aid; test breaks
McCormack	Oral delivery, live/in-person
Noakes et al.	Speech recognition system
Quesen & Lane	Extended time; oral delivery, live/in-person; small group; test breaks
Robinson et al.	Student reads aloud (to self)
Wise et al.	Prompting (“proctor notification”)

Table E-2. Presentation Accommodations Itemized by Study

Author/s	Electronic administration	Large print/magnification	Oral delivery, live/in-person	Oral delivery, not specified	Prompting	Student reads aloud (to self)	N
Castro et al.	•						1
Giusto & Ehri			•		•		2
Hahn et al.	•						1
Kern et al.		•		•			2
McCormack			•				1
Quesen & Lane			•				1
Robinson et al.						•	1
Wise et al.					•		1
TOTAL studies (of 8)	2	1	3	1	2	1	

Table E-3. Equipment Accommodations Itemized by Study

Author/s	Electronic administration	N
Castro et al.	•	1
Hahn et al.	•	1
TOTAL studies (of 2)	2	

Table E-4. Response Accommodations Itemized by Study

Author/s	Calculator	Electronic admin- istration	Dictated response	Speech recogni- tion system	N
Castro et al.		•			1
Hahn et al.		•			1
Kern et al.	•		•		2
Noakes et al.				•	1
TOTAL studies (of 4)	1	2	1	1	

Table E-5. Scheduling Accommodations Itemized by Study

Author/s	Extended time	Test breaks	N
Kern et al.	•		1
Quesen & Lane	•	•	2
TOTAL studies (of 2)	2	1	

Table E-6. Setting Accommodations Itemized by Study

Author/s	Individual	Small group	Specialized setting	N
Kern et al.	•	•	•	3
Quesen & Lane		•		1
TOTAL studies (of 2)	1	2	1	

Appendix F

Findings for K–12 Studies in 2019

Authors	Findings Statement	Effects	Implement/Use	Perceptions	Test Items	Validity	Content
Buzick	<p>The researcher examined the consistency of assignment of accommodations to students with disabilities in grades 3–8 across two consecutive years in two states. Most students were assigned accommodations in both years, yet approximately 25% were not. The second most frequent pattern of accommodations assignment was not receiving accommodations in the first year but receiving them in the second year. The researcher found correlations between inconsistent use of accommodations across academic years and different rates of learning growth in English language arts and math. Students assigned accommodations only in the second year averaged the largest academic performance improvement compared to students with other patterns of accommodations. These growth rate differences were found when using two different measurement approaches. The implications of these findings were discussed in consideration of assignment consistency, inferences about growth scores, and whether a systematic policy change might threaten validity of score interpretation.</p>		X				Math, Other LA
Carey et al.	<p>Surveyed educators demonstrated positive beliefs, knowledge, and practices regarding testing accommodations. Scores on the beliefs domain were significantly higher than those for practice and knowledge. Beliefs and knowledge were found to be significantly associated with practice. High school special education teachers had the most positive beliefs about accommodations, while general education teachers held the lowest (yet still positive) beliefs. No significant differences were found between elementary and middle school special and general education teachers. Teachers overall, especially general education teachers, reported a lack of confidence in their knowledge and ability to appropriately provide accommodations and in their ability to instruct students on accommodation use. Differences in knowledge ratings existed depending on the training teachers had received on differentiated instruction and universal design for learning. Special education teachers supported practices such as modeling and discussing student accommodations for students to benefit from their accommodations. In contrast, general education teachers tended not to do so, especially at the high school level.</p>		X	X			NA

Authors	Findings Statement	Effects	Implement/ Use	Perceptions	Test Items	Validity	Content
Castro et al.	<p>Elementary students in three schools performed similarly on average on a cognitive skills assessment (i.e., Wechsler Intelligence Scale for Children) when responding to a paper-and-pencil version and a version administered on an electronic tablet. This finding existed for the Syndrome scales (e.g., withdrawn/depressed; language/thought problems; oppositional; attention problems) and for the Broad scales (internalizing, externalizing, total problems). Comparable groups of students who were identified with various difficulties during testing—including depression, anxiety, behavioral disabilities, attention problems, and 'language/thought problems'—had no significant mean score differences between the two testing conditions. The authors conclude that the results are important for practitioners because they indicate that the format of a test does not increase the likelihood of emotional or behavioral problems during testing.</p>	X					Cognitive Skills
Giusto & Ehri	<p>Students with decoding-related reading disabilities scored significantly higher on reading comprehension when provided by test proctors with both prompting ("pacing") support and partial oral delivery—of test instructions and item stems but not reading passages—than with either the pacing-only support or without accommodations. Students using only prompting support did not score significantly differently than students not using accommodations. The students without decoding difficulties (i.e., without disabilities) did not score significantly differently across the three testing conditions. Additionally, the varying linguistic backgrounds among students in both participant groups did not demonstrate an effect on the mean scoring patterns; that is, linguistic background was not a factor in the differences reported previously. The researchers noted that the partial read-aloud accommodation with pacing support can address concerns indicated in other research and practice about invalidating the reading comprehension construct. In other words, not reading the text passages yet providing read-aloud for the remaining parts of the test was demonstrated to differentially benefit students with decoding skills difficulties.</p>	X				X	R

Authors	Findings Statement	Effects	Implement/Use	Perceptions	Test Items	Validity	Content
Hahn et al.	<p>The researchers determined that the effects of using a multisensory tablet—with a touchscreen employing haptic features, providing both sound and vibration feedback—were similar to using embossed paper tactile graphic images on mathematics and science assessments for individuals with visual impairments. Paired sample mean score comparisons yielded that participants had 6% more correct answers when getting test item information through embossed graphics on paper than through the multisensory tablet; this difference was not statistically significant. Factors such as the severity of participants' visual impairments, and their experience with braille reading, and even not having previous experience with the tablet did not mediate these effects. Participants' perceptions of using the multisensory tablet included both positive and negative feedback. Positive feedback indicated that users enjoyed using the device and that the sounds and vibrations were engaging, while negative feedback indicated concern of line thickness, use of space, and repetitive sounds. Researchers reported that this feedback could be incorporated easily to improve images.</p>	X		X			M, S
Kern et al.	<p>Authors conducted a descriptive analysis of the proportions of 222 students receiving various supports, by category and specific accommodation, in the classroom and during state or district assessments. [Note: We do not report use data for instructional accommodations and homework supports here.] Accommodations provided on classroom tests totaled 304, including the test-related accommodations of extended test time-48%, small group testing-48%, and oral delivery-41%. State/district assessment accommodations totaled 685: Presentation-55%; Response-27%, Timing/Scheduling-58%; Setting-82%. Most students received more than one accommodation. They included extended time-56%; small group-90%; oral delivery-52%. The number and category of accommodations provided to students with disabilities tended not to be associated with race/ethnicity, gender, or grade level. Few links were found between number and type of accommodations and participants' academic, behavioral, or emotional functioning. For both classroom and state assessments, students with lower math performance (on WJ-III) were more likely to be assigned response accommodations, including calculators. During state assessments, students with lower math and reading (on WJ-III) had more likelihood of presentation accommodations, and students with higher anxiety (per MASC scores) received more checking by proctor, such as checking for understanding.</p>		X				M, R

Authors	Findings Statement	Effects	Implement/ Use	Perceptions	Test Items	Validity	Content
McCormack	<p>Analyses of two years of New York grade 4 ELA test data showed a significant difference in the use of “tests read” (oral delivery, live/in-person) across a sample of the 37 educational service regions of the state. The sample regions excluded those with too few students with disabilities, those that did not use the accommodation, and New York City. The proficiency levels of 3,244 students with disabilities were significantly different among regions, but post hoc tests did not indicate that the live/in-person oral delivery accommodation had an impact on scores. Results of a survey of 37 special education administrators representing all regions of the state indicated that the oral delivery accommodation was assigned to “level the playing field” rather than to provide access. The survey also revealed several issues, including a lack of knowledge about or use of the state’s accommodations guidelines or decision-making tool. Similarly, survey results suggested there was a lack of training for teachers on accommodations in general, and specifically on the state’s decision-making tool.</p>		X				Other LA
Noakes et al.	<p>Three participants with traumatic brain injuries performed better when using speech-to-text assistive technology than when handwriting their response to narrative writing prompts in total words written (TWW), words spelled correctly (WSC), and correct writing sequences (CWS). Participants’ TWW scores in the handwriting condition averaged from 2 to 23, compared to speech-to-text scores which ranged from 44 to 101. Comparisons showed large effect sizes for all three students, favoring the assistive technology condition. For WSC, students’ handwriting scores averaged from 0 and 23 while speech-to-text scores averaged from 44 and 101. For CWS, students’ handwriting scores averaged from 0 to 23 compared to their speech-to-text CWS scores, which averaged from 42 to 87. On all measures, scores in the two conditions were non-overlapping with large effect sizes for the three students with TBI, always favoring the speech-to-text condition. Participants’ handwritten products were described as oversized and either unrecognizable letters and words, or under-developed lettering (in comparison with peers without TBI).</p>	X				W	
Quesen & Lane	<p>Regardless of the four differential item functioning (DIF) analysis methods—logistic regression; hierarchical generalized linear model (HGLM), Wald-1 IRT-based test; and Mantel-Haenszel—applied to the 60 multiple-choice items on a statewide data set of grade 8 math test scores, no DIF was detected when the reference group (non-accommodated students without disabilities) had a similar performance distribution to the focal group (accommodated students with disabilities). Each of the methods detected DIF when the reference group was not similar in ability to the focal group. The authors concluded that creating the reference group to be similar in performance to the focal group can help avoid erroneous DIF detection for students with disabilities.</p>				X	M	

Authors	Findings Statement	Effects	Implement/Use	Perceptions	Test Items	Validity	Content
Robinson et al.	Students with reading-related disabilities in grades 2 and 3 benefited differentially more on retell performance—a measure of reading comprehension—when reading aloud to themselves, compared to reading silently to themselves. In contrast, students with reading disabilities in grades 4 and 5 had similar performance when using these different reading modalities. All student participants demonstrated improvement in comprehension across the school year, from fall to spring, for both reading modalities. The researchers concluded that an underlying factor was that students early in their reading development can benefit particularly well from reading aloud to themselves, a phenomenon that seems to fade for students with higher reading skill development.	X					R
Wise et al.	Higher test-taking engagement and higher test performance resulted when test engagement monitoring (i.e., effort monitoring) and proctor notification were used. These effects were stronger for computer-based reading tests than for computer-based math tests. In addition, correlations between beginning and final test scores were higher when test engagement monitoring and proctor notification were used than when they were not used, supporting convergent validation. Results from a sample of test proctors suggested that, in general, test proctors viewed the proctor notification feature as valuable.	X					M, R
TOTAL		6	4	2	1	1	

KEY for Appendix F

Effects	Compare effects of accommodations on assessment scores
Implementation/Use	Report on implementation practices and accommodations use
Perceptions	Study/compare perceptions and preferences about use
Test Items	Compare test items across assessment formats
Content	Academic content area
M	Mathematics
R	Reading
Other LA	English language arts (other than reading)
S	Science
W	Writing
Cognitive skills	Norm-referenced cognitive skills assessment (i.e., IQ test)

INSTITUTE *on* COMMUNITY INTEGRATION

UNIVERSITY OF MINNESOTA

NCEO is an affiliated center of the Institute on Community Integration