Hammill Institute on Disabilities

Journal of Emotional and Behavioral Disorders
2019, Vol. 27(3) 178–192

DOI: 10.1177/1063426618763108
jebd.sagepub.com
SAGE

An Evaluation of IEP Accommodations for Secondary Students With Emotional and Behavioral Problems

Lee Kern, PhD[1], Allyse A. Hetrick, MEd[1], Beth A. Custer, PhD[1], and Colleen E. Commisso, MS[1]

Abstract

Accommodations are intended to address student academic and behavioral deficits by reducing obstacles that impede learning and accurately measuring skills. There is limited research, however, pertaining to the types of accommodations students receive and their selection, particularly among those with emotional and behavioral problems. This is a significant concern for secondary age students who spend the majority of their day in regular education settings and must participate in high-stakes testing. We examined types of accommodations provided to 222 secondary students with emotional and behavioral problems, their use (i.e., classroom or standardized assessments), and variables related to their selection. Analyses indicated (a) students received a wide array of accommodations with some differences depending on disability type, (b) more accommodations were provided in the classroom than on standardized testing, (c) few demographic variables were associated with type or number of accommodations, and (d) with a single exception, academic and behavioral functioning did not explain type of accommodation received. The findings suggest that accommodation selection is highly imprecise and point to the critical need for further research in this area.

Keywords

students, emotional and behavior problems, accommodations, high school

Students with emotional and behavioral problems have historically had poor outcomes, both behaviorally and academically (Newman et al., 2011; Simpson, Peterson, & Smith, 2011). For instance, the vast majority of students with emotional and behavioral disorders (EBD) performs well below grade level in literacy and math and receives more failing grades than students in any other disability group (Bradley, Doolittle, & Bartolotta, 2008). This poor school performance also is mirrored on standardized assessments, with research indicating that fewer than 50% meet the minimum standard in reading, 80% score basic or below basic on measures of writing, and fewer than 34% meet proficiency standards in math (Carr-George, Vannest, Willson, & Davis, 2009; Gage, Wilson, & MacSuga-Gage, 2014; Temple-Harvey & Vannest, 2012). Clearly, this population of students struggles to perform adequately in the classroom and on high-stakes assessments.

One support intentionally designed to help students overcome performance obstacles that both impede learning and mask the accurate measure of skills during testing situations is accommodations. The term *accommodation* has been defined as a change in the way an assignment or test is administered to "level the playing field." Accommodations should allow students with disabilities to demonstrate their knowledge without affecting the validity of the assignment or test results. That is, they should not reduce or change the expectations of the assignment or test, but instead should remove barriers resulting from the interaction of the student's disability and the assignment/test characteristics (Fuchs, Fuchs, & Capizzi, 2005; Fuchs, Fuchs, Eaton, Hamlett, & Karns, 2000; Ketterlin-Geller, Yovanoff, & Tindal, 2007; Lang, Elliott, Bolt, & Kratochwill, 2008). A valid accommodation does not change the construct being evaluated but offers a "differential boost" to students with disabilities, improving their performance to a greater extent compared with the performance of their nondisabled peers, hence compensating for their disability (Elliott & Marquart, 2004; Feldman, Kim, & Elliott, 2011). Research indicates that an appropriate accommodation can result in a differential boost for students with disabilities by 0.26 to 0.41 standard deviations (Fuchs & Fuchs, 2001; Lang et al., 2008).

[1]Lehigh University, Bethlehem, PA, USA

Corresponding Author:
Lee Kern, College of Education, Lehigh University, 111 Research Drive, Bethlehem, PA 18015, USA.
Email: lek6@lehigh.edu

Identifying appropriate accommodations has become particularly important for several reasons. First, high-stakes tests are increasingly utilized for important decisions, such as course placement, grade promotion, and high school graduation (Christenson, Decker, Triezenberg, Ysseldyke, & Reschly, 2007; Ysseldyke et al., 2004). For example, 23 states currently require that students pass proficiency exams to graduate from high school. Second, a large percentage of students spend time in the general education setting. For example, Wagner and colleagues (2006) found that 85.4% of high school students with EBD spend more than half of their day in the general education classroom. More recent data from the National Center for Education Statistics (U.S. Department of Education, National Center for Education Statistics, 2016), although not disaggregated by age group, also indicate a large percentage of students with EBD are placed in general education. Specifically, 82.7% of students with EBD, aged 6 to 21, spend at least part of their day in the general education classrooms. Supports are needed for these students to fully benefit from instruction in the general education environment. Finally, Individuals With Disabilities Education Act (IDEA; 2004) stipulates that eligible students receive accommodations (a) to enable them to make progress in the general education curriculum and (b) that are necessary to measure academic achievement on state/district-wide assessments. Both types of accommodations are required components of the Individualized Education Program (IEP), as described in the federal regulations (34 C.F.R. §300.320[a]).

Although there are many potential accommodations, we reviewed the most recent online policy document from each state (e.g., http://www.education.pa.gov/Documents/K-12/Assessment%20and%20Accountability/PSSA/2017%20Accommodations%20Guidelines.pdf) describing accommodations. These guidelines indicated that the majority (including D.C.) groups accommodations into four categories. Specifically, 36 describe accommodations in the following four areas: Presentation, Response, Timing/Scheduling, and Setting. The remaining states either list accommodations without categorizing ($n = 5$) or use categorical variations, such as fewer categories or the inclusion of an "other" category ($n = 10$). For the purpose of the current study, we aligned our coding categories with the preponderance of states.

Accommodations in the Presentation category adjust the way information is given or presented (e.g., larger print, read aloud). Response accommodations allow students to demonstrate knowledge in different ways (e.g., using a word processor instead of writing by hand, using a calculator). Timing/Scheduling accommodations allow flexibility in the timing of an assessment (e.g., providing breaks, extended time). Finally, Setting accommodations refer to where in space the assessment occurs (e.g., in a small group, in a specified seat near the teacher).

IEP team members are responsible for determining student characteristics or deficits that impede performance and selecting accommodations that will compensate for those deficits. This applies both to classroom instruction and assessments. However, despite extensive delineation of accommodations in state legislation and practitioner resources (e.g., books, guidelines), limited research is available to guide their selection, particularly for students with emotional and behavioral problems. For example, Harrison, Bunford, Evans, and Owens (2013) conducted a systematic review of literature and identified 18 peer-reviewed studies evaluating 12 potential accommodations for students with EBD and attention deficit hyperactivity disorder (ADHD). They concluded that few of the identified strategies could be classified as accommodations and there was very little evidence supporting their effectiveness for students with EBD and ADHD, highlighting the need for further research.

Although effective practice can foreshadow research findings, this does not appear to be the case with accommodations. Studies show that instead of considering student characteristics when selecting accommodations, teachers rely on variables such as their opinion of fairness or ease of administration (Maccini & Gagnon, 2006; McKevitt & Elliott, 2003; Schulte, Elliott, & Kratochwill, 2001; Ysseldyke et al., 2001). Moreover, when student characteristics are considered, educators often focus on those that are irrelevant. For instance, research indicates that race, free and reduced lunch status, reading level, and previous retention are predictive of receiving accommodations (Fuchs & Fuchs, 2001; Fuchs, Fuchs, Eaton, Hamlett, Binkley, & Crouch, 2000). In addition, studies indicate that teachers have difficulty differentiating between students who would or would not benefit from testing accommodations, resulting in the overuse of accommodations that are not beneficial (Fuchs, Fuchs, Eaton, Hamlett, Binkley, & Crouch, 2000; Fuchs, Fuchs, Eaton, Hamlett, & Karns, 2000; Helwig & Tindal, 2003). Research also documents that secondary general education teachers are less willing to provide testing accommodations to students with EBD compared with special education students with other disabilities (Maccini & Gagnon, 2006).

Together, the aforementioned research indicates that accommodation selection is challenging for teachers. Procedures that teachers use to select accommodations (a) are not matched to student need and therefore may not be optimally beneficial; (b) are often selected based on irrelevant student characteristics; and (c) may not be provided impartially across disability groups. Furthermore, the limited existing research has been conducted almost exclusively with younger participants at the elementary grade level, a concern echoed by Harrison et al. (2013). Although a few studies have examined middle school participants, they have not included students higher than Grade 8. In addition,

almost all accommodations studies have examined academic accommodations implemented with students with specific learning disabilities (SLD), with few studies examining behavioral accommodations (Feldman et al., 2011). Finally, research has not yet examined accommodations used both in the classroom and for standardized tests. Thus, the general purpose of this study was to investigate the types of classroom and testing accommodations received by high school students with emotional and behavioral problems. The following specific research questions were addressed:

Research Question 1: What types of accommodations do high school students who exhibit emotional and behavioral problems receive in the classroom and on state/district-wide assessments? Does accommodation type differ depending on special education classification?

Research Question 2: Do types of accommodations that high school students with emotional and behavioral problems receive in the classroom and on state/district-wide assessments differ depending on student demographic variables (i.e., gender, disability category, race, grade level, or state of residence)?

Research Question 3: Does the number of accommodations that high school students with emotional and behavioral problems receive in the classroom and on state/district-wide assessments differ depending on student demographic variables (i.e., gender, disability category, race, grade level, or state of residence)?

Research Question 4: Among high school students identified as having emotional and behavioral problems, do the types of accommodations received, either in the classroom or on state/district-wide assessments, differ depending on their academic, emotional, or behavioral functioning?

Method

Participants

Data from a larger study (Center for Adolescent Research in Schools [CARS]) were used to answer the research questions. CARS was a national center funded by the Institute of Education Sciences (IES) with the purpose of developing and evaluating a multicomponent intervention package designed to improve outcomes for high school students with severe social, emotional, and behavioral problems (Kern et al., 2015). The package was evaluated using a 2-year randomized controlled trial (RCT).

Fifty-four high schools across five states participated in the RCT. Schools were selected in Kansas (n = 5), Missouri (n = 7), Ohio (n = 16), Pennsylvania (n = 10), and South Carolina (n = 16) based on proximity to the universities of study researchers and school staff willingness to engage in project activities. Participating schools were fairly evenly distributed with respect to community location (defined by the U.S. Department of Education), with 21 (39%) suburban, 20 (37%) rural, and 13 (24%) urban.

A school liaison (typically a school counselor, administrator, or special education teacher) referred participants who met the following study criteria for participation: (a) would be attending ninth to 11th grade during Year 1 of the study (i.e., 2011–2012 academic year) and (b) exhibited serious social, emotional, and/or behavioral problems. Students were screened by CARS staff to assure significant social, emotional, or behavioral impairment, indicated by (a) a *T*-score of 60 or higher on either the internalizing or externalizing composites of the *Behavior Assessment System for Children, Second Edition–Teacher or Parent Version* (BASC-2; C. R. Reynolds & Kamphaus, 2004); (b) a *T*-score of 60 or higher on the *Multidimensional Anxiety Scale for Children* (MASC; March, 1998); or (c) a *T*-score of 60 or higher on the *Reynolds Adolescent Depression Scale, Second Edition* (RADS-2; W. M. Reynolds, 2002).

In addition, students also needed to demonstrate impairment in school functioning by exhibiting any two of the following: (a) four or more office discipline referrals/behavioral infractions across the semester prior to enrollment or five or more in any month of the semester screened, (b) five or more absences (other than illness) or tardies to class in any month of the semester screened or previous semester, (c) two or more in- or out-of-school suspensions in the academic year screened or previous academic year, or (d) one or more Fs or two or more Ds in any core academic subject in one of two grading periods most recent to screening. Performance during previous semesters was considered because screening began during the summer.

Students with autism spectrum disorder (ASD) were excluded. In addition, students with an IQ score below 75 were also excluded to assure understanding of concepts in some of the interventions (e.g., cognitive behavior therapy). Finally, students had to have at least one parent/guardian who could speak English fluently to complete assessments.

A total of 647 participants met eligibility criteria and agreed to participate. Of the total sample, 49% (n = 317) had a special education label, whereas the remaining 51% of students (n = 330) had no label. IEP data were analyzed for students who had a special education label.

In spite of requests, circumstances prevented collection of all IEPs from all special education students (i.e., some schools would not provide IEPs in spite of parent consent, students moved before they could be obtained, no IEP existed). Therefore, IEP data were available for 222 participants, representing 70% of special education students. The sample was 73% male (n = 162) and 27% female (n = 60). Primary disability categories included 50.5% identified as SLD (n = 112), 25.2% as EBD (n = 56), 20.7% as other health impairment (OHI; n = 46), and 3.2% as other (n = 7; for example, traumatic brain injury [TBI], speech impairment). Special

Table 1. Participant Characteristics.

Characteristics	Total sample	Disability category			
		SLD	EBD	OHI	Other
Total sample	222 (100%)	112 (50.5%)	56 (25.2%)	46 (20.7%)	7 (3.2%)
Gender					
Male	162 (73%)	81 (72.3%)	44 (78.6%)	30 (65.2%)	6 (85.7%)
Female	60 (27%)	31 (27.7%)	12 (21.4%)	16 (34.8%)	1 (14.3%)
Race					
White	119 (53.6%)	61 (54.5%)	32 (28.6%)	23 (50%)	2 (28.6%)
Non-White	103 (46.4%)	51 (45.5%)	24 (21.4%)	23 (50%)	5 (71.4%)
Grade					
8	6 (2.7%)	3 (2.7%)	1 (1.8%)	1 (2.2%)	1 (14.3%)
9	62 (27.9%)	30 (26.8%)	19 (33.9%)	12 (26.1%)	1 (14.3%)
10	103 (46.4%)	55 (49.1%)	22 (39.3%)	23 (50%)	2 (28.6%)
11	44 (19.8%)	21 (18.8%)	12 (21.4%)	8 (14.3%)	3 (42.9%)
State of residence					
Pennsylvania	51 (23%)	27 (24.1%)	18 (32.1%)	5 (10.9%)	1 (14.3%)
Ohio	66 (29.7%)	35 (31.3%)	18 (32.1%)	9 (19.6%)	3 (42.9%)
Missouri	18 (8.1%)	8 (7.1%)	2 (3.6%)	7 (15.2%)	1 (14.3%)
South Carolina	64 (28.8%)	34 (30.4%)	14 (25%)	14 (30.4%)	2 (28.6%)
Kansas	23 (10.4%)	8 (7.1%)	4 (7.1%)	11 (23.9%)	0 (0%)

Note. SLD = specific learning disability; EBD = emotional and behavioral disorder; OHI = other health impairment.

education label was missing for one student. Table 1 shows sample demographic characteristics.

Measures

BASC-2. The BASC-2 is a broad assessment of a child's emotional and behavioral functioning. A parent/legal guardian completed the parent rating form (150 items; C. R. Reynolds & Kamphaus, 2004). Behaviors are rated on a 4-point scale, including 1 (*never*), 2 (*sometimes*), 3 (*often*), and 4 (*almost always*). The standard scores of the externalizing composite were used as a measure of the students' behavior outcomes as reported by parents. *T*-scores of 50 represent an average score, with higher scores indicating greater levels of problem behavior. *T*-scores of 60 or above generally indicate students are "at risk" of developing clinically significant problems, whereas *T*-scores of 70 or above indicate clinical significance. The assessment is suitable and normed for high school students. The BASC-2 has good psychometric properties with internal consistency ranging from .80 to .90, test–retest reliability of .82 across age ranges, long-term stability of .69, and convergent validity at $r = .81$.

MASC. The MASC is a 39-item self-report assessment of anxiety-related symptoms in youth 8 to 18 years old (March, 1998). It assesses a broad range of emotional, physical, cognitive, and behavioral symptoms that represent dimensions of childhood anxiety. The scale provides four main scores for social anxiety, separation anxiety, harm avoidance, and physical symptoms, as well as a total score. Students rate their own behavior on a 4-point Likert-type scale with scores and anchors 0 (*never true about me*), 1 (*rarely true about me*), 2 (*sometimes true about me*), and 3 (*often true about me*). *T*-scores of 65 or above generally indicate level of symptoms associated with clinical anxiety. The measure has good reported psychometric properties with alpha coefficients from .87 to .89 and test–retest reliability from .73 to .89 (March, Sullivan, & Parker, 1999; Thaler, Kazemi, & Wood, 2010).

RADS-2. The purpose of this 30-item self-report assessment is to identify depressive symptoms in adolescents ranging in age from 11 to 20 years (W. M. Reynolds, 2002). It measures the four basic dimensions of depression: dysphoric mood, negative affect, negative self-evaluation, and somatic complaints. Students choose response options arranged on a 4-point Likert-type scale: 1 (*almost never*), 2 (*hardly ever*), 3 (*sometimes*), and 4 (*most of the time*). The RADS-2 standard score provides an indication of the clinical severity of an individual's depressive symptoms. *T*-scores of 60 or above indicate level of symptoms associated with clinical depression. The scale is widely used and has good reported overall psychometric properties with internal consistency ranging from .92 to .94 and test–retest reliability at .89 (W. M. Reynolds, 2002).

Woodcock Johnson Tests of Achievement, Third Edition (WJ-III). The WJ-III is a battery of tests to assess student

achievement in reading, writing, and mathematics. The Broad Reading standard score (i.e., Letter–Word Identification, Reading Fluency, Passage Comprehension subtests), and the Broad Math standard score (i.e., Calculation, Math Fluency, and Applied Problems subtests) were used to measure student academic achievement (Woodcock, McGrew, & Mather, 2001). Overall, the WJ-III has strong psychometric properties, with an internal consistency reliability of .98 and an interrater reliability range of .93 to .99. Stability scores indicate that the WJ-III is a highly reliable test with correlations ranging from .85 to .96. In terms of validity, the WJ-III correlates moderately with *Wechsler Individual Achievement Test* (r = .65) and with *Kaufman's Test of Educational Achievement, Second Edition* (r = .79).

Demographic data. Parents completed a demographic questionnaire about their child and family prior to the start of the RCT. Demographic data analyzed in the current study were gender (i.e., male, female), race (i.e., White, non-White), grade level (i.e., 8, 9, 10, 11), primary disability category (i.e., SLD, EBD, OHI, Other), and student state of residence (i.e., Pennsylvania, Ohio, Missouri, South Carolina, Kansas).

Assessment and Data Collection Procedures

Assessments were administered at several time points throughout the CARS RCT for students in both the treatment and control groups. All assessments were individually administered to students by trained project staff, either in the home or at school. Assessments were completed using teleforms that were sent to the Texas Institute for Measurement, Evaluation, and Statistics (TIMES) at the University of Houston for entry, storage, and analysis. In addition, CARS facilitators requested each of the participating students' IEPs from the schools during each year of intervention. For the current study, data were used from measures administered at baseline, and the earliest IEP available after consent was obtained.

Coding Categories

Each accommodation was coded into one of four categories: (a) Presentation, (b) Response, (c) Timing/Scheduling, and (d) Setting. Although most states provide examples of allowable accommodations, few provide operational definitions. Therefore, we relied on definitions used in the literature (Christensen, Braam, Scullin, & Thurlow, 2011). Presentation accommodations were defined as altering material(s) or test(s), so it is presented in a nonstandard manner. Examples included presenting material in an auditory or multisensory format (e.g., read aloud); amplifying teacher talk; providing a study guide, checklist, or teacher notes; altering print or format; simplifying directions*; providing concrete examples prior to task*; or offering hands-on materials or manipulatives*. Response accommodations were defined as any change in the way a student responded to questions, assignments, or activities. Examples included using electronic devices (e.g., speech to text translator, calculators) or organizers when responding, dictating to a scribe, providing graph paper to align responses, or highlighting answers (rather than writing out). Timing/Scheduling accommodations were defined as changing the standard timing of the activity, assignment, or assessment. Examples included providing extended time, allowing breaks from setting or activity, chunking long-term assignments into smaller parts, or allowing access to staff (i.e., nurse, guidance, case manager) during an assignment or assessment in response to emotional/behavioral need (e.g., expression of anxiety or frustration, behavioral escalation)*. Setting accommodations were defined as changes in the usual classroom location or structure where a student received instruction or participated in an assessment. Examples included providing preferential seating (e.g., increasing proximity to teacher, seating in front of classroom) or completing assessments/assignments in another setting (e.g., special education classroom, small group, 1:1 context). Accommodations that did not fall into one of the above four categories were coded as "Other."

After coding was completed, a large number of accommodations (n = 370) fell into the Other category. The authors reviewed and discussed these accommodations to identify similar themes. A few of the accommodations aligned with the four previously defined categories, but had not been included as examples in previous definitions (Christensen et al., 2011). Thus, we expanded examples in two of the categories (Presentation and Timing/Scheduling), noted by an asterisk in the definition above. In addition, the following three new categories emerged: (a) Check, (b) Cues/Prompts, and (c) Structured Behavioral Strategies. Although these categories have not been previously described in the literature, they reoccurred with a relatively high frequency across IEPs and therefore emerged as new categories.

Check accommodations were defined as "checking in" with the student to increase the likelihood that he or she understood the content, directions, or behavioral expectations. Examples included asking a student about how he or she should behave in a novel situation (setting, new task), periodically monitoring student progress on an assignment or assessment to ensure understanding, asking the student to repeat directions, and assisting the student with organization (e.g., checking to assure homework was recorded). Cue/Prompt accommodations were defined as providing behaviorally related verbal or visual prompts, cues, or redirections, as needed, when it was evident that the student should have had previous knowledge of the behavioral expectations. Examples included reminding student of

rules, providing cues/prompts (including electronic) of expected behavior, providing verbal or gestural redirection. Structured Behavioral Strategies were defined as systematic approaches such as positive behavior support plans, consistent use of positive reinforcement for following expectations, daily behavioral report cards or point sheets, and/or delivering specific praise for appropriate behavior.

Coding Procedures

Consistent with federal IDEA requirements (34 C.F.R. § 300.320[a][4]), each state's IEP included a section for classroom accommodations, modifications, or specially designed instruction that would enable the student to make progress in the general education curriculum. In addition, a separate section in each IEP described the accommodations on state/district-wide assessments (§ 300.320[a][6]). These two sections allowed us to identify accommodations and code how they were applied.

Four graduate students in special education, counseling psychology, or school psychology served as coders and were trained during an initial meeting in which category definitions and examples were discussed. Prior to coding into categories, accommodations were parceled out from modifications, specially designed instruction and descriptions of other changes/events. Among the 222 IEPs reviewed, there were 1,925 statements describing accommodations, modifications, specially designed instruction, or other changes.

For this preliminary sorting, accommodations were defined as any alteration to a test or assignment format or the administration procedures that did not reduce the content knowledge expectation of the assignment or test. We excluded modifications, defined as changes to practices in schools that alter, lower, or reduce expectations to compensate for a disability, such as providing text and assignments at an instructional level lower than the student's current grade level (Harrison et al., 2013). We also excluded specially designed instruction (inconsistent with an accommodation), defined in IDEA as adaptations to

> the content, methodology, or delivery of instruction (i) to address the unique needs of the child that result from the child's disability; and (ii) to ensure access of the child to the general curriculum, so that the child can meet the educational standards within the jurisdiction of the public agency that apply to all children. (§ 300.39[b][3])

Although this federal definition could conceptually include accommodations, we excluded specially designed instruction that provided supplemental instruction and/or specific skill instruction (e.g., instruction in organizational skills, direct instruction in social skills, direct instruction in reading fluency). After excluding 60 descriptions of modifications and specially designed instruction, 1,865 accommodations remained. However, an additional 25 descriptions were excluded because they could not be considered an accommodation, modification, or specially designed instruction, or were too vague to code (e.g., "random and frequent searches," "additional support in mainstreamed class," "physical restraint used if necessary," "resource assistance"). Thus, 1,840 accommodations remained for coding.

To facilitate coding, the coders first recorded, verbatim, wording from each individual accommodation into an Excel spreadsheet. For some IEPs, wording was modified slightly for consistency purposes (e.g., "allow student to use calculator" was changed to "calculator"; "test in small group" was changed to "small group testing"). Note also that some accommodations listed under "classroom accommodations, modifications, and specially designed instruction" (e.g., read aloud, extended time) specified application to assignments and/or tests. Others, however (i.e., read aloud, extended time, small group), did not indicate whether they applied to assignments or tests and are reported as "not specified" in results. Coders then practiced independently categorizing accommodations listed on several IEPs into the initial four and Other coding categories and agreements and disagreements were discussed as a group. This process required extensive discussion because language in IEPs was often vague. The aforementioned process continued until five IEPs were independently coded with accuracy exceeding 90%.

Subsequently, accommodations in the Other category were recoded using the three new categories (while still maintaining an Other category). As with the initial coding, definitions and examples for each category were discussed. In the case of disagreements, the first author made a final determination regarding in which category an accommodation belonged.

Intercoder Agreement

After the initial coding with the first four categories and Other, 55 of the 222 IEPs (25%) were randomly selected to assess intercoder agreement. The 55 IEPs contained 457 individual accommodations, representing 24% of the 1,925 total accommodations coded. Initial coders coded IEPs they did not previously code. Agreement was assessed on a point-by-point basis for each individual accommodation whereby a category match was considered an agreement. Overall agreement was calculated by dividing the number of category agreements by the number of category agreements plus disagreements. Total agreement during the initial coding was 97%.

After recoding accommodations in the Other category, intercoder agreement was assessed for 35% of the accommodations (130 of the 370) in the three new categories (Check, Cues/Prompts, Structured Behavioral Strategies).

Agreement was calculated in the same manner as the initial coding (exact category agreements divided by agreements plus disagreements). Total agreement for the second coding was 88%.

Data Analysis

The first research question examined what types of accommodations high school students with emotional and behavioral problems received in the classroom and on state/district-wide assessments. To answer this question, data were analyzed descriptively by determining the overall number and percentage of students who received each type of accommodation. In addition, the number and percentage of students who received each type of accommodation was analyzed for each primary disability category group (i.e., SLD, EBD, OHI, Other).

The second research question investigated relationships between accommodation types and student demographic characteristics among high school students with emotional and behavioral problems. To answer this question, Pearson's chi-square tests were conducted for each type of accommodation (i.e., Presentation, Response, Timing/Scheduling, Setting, Check, Cues/Prompts, Structured Behavioral Interventions) to determine whether there was a relationship to demographic characteristics (i.e., gender [male, female], grade level [8, 9, 10, 11], disability category [SLD, EBD, OHI, Other], race [White, non-White], or state of residence [Pennsylvania, Ohio, Missouri, Kansas, South Carolina]). The accommodation type variable was dichotomous, indicating whether a student did or did not receive at least one accommodation within each specific category (i.e., 0 = student received zero accommodations in the category, 1 = student received at least one accommodation within the category). The chi-square test has two important assumptions relating to independence and expected frequencies. First, every observation should fall into one and only one category (cell). In addition, all expected counts should be greater than 1 and no more than 20% of the cells should have an expected frequency of less than five (Field, 2013). These criteria were met for all analyses. Effect sizes were evaluated using Cramer's V, a comparative statistic that shows the strength of the observed relationships between two variables (Cramer, 1999). Furthermore, to determine which cells contributed the most to significant associations, adjusted standardized residuals were observed (Agresti, 2013). According to Agresti (2013), adjusted standardized residuals are similar to z scores in that they are normally distributed with a mean of zero and a standard deviation of one. Thus, adjusted residual values that exceed 2 in absolute value indicate that the frequency of a cell contributes significantly to the association. Accommodations used in the classroom and those designated for state/district-wide assessments were analyzed separately.

The third research question examined differences in the number of accommodations that high school students with emotional and behavioral problems received depending on student demographic variables. Because the total number of accommodations that students received was not normally distributed, nonparametric tests were used to answer this question. For the two-group demographic variables of gender (i.e., male, female) and race (i.e., White, non-White), the Mann–Whitney U Test was used to examine differences between the independent demographic groups and the total number of accommodations students received. Similar to the Mann–Whitney U Test, the nonparametric Kruskal–Wallis Test allows comparison of three or more groups. This test was used to compare the total number of accommodations among the demographic variables of grade level (i.e., 8, 9, 10, 11), disability category (SLD, EBD, OHI, Other), and state of residence (Pennsylvania, Ohio, Missouri, Kansas, South Carolina). As a follow-up to significant Kruskal–Wallis Tests, Mann–Whitney Tests between pairs of groups were analyzed. To control for Type 1 errors, a Bonferroni adjustment was applied. Effect size values of r were calculated. Accommodations used in the classroom and those designated for state/district-wide assessments were analyzed separately.

The fourth research question explored differences in the types of accommodation that high school students with emotional and behavioral problems received in the classroom and on state/district-wide assessments depending on their academic, emotional, or behavioral functioning. To address this question, a series of independent samples t tests were conducted comparing the independent variable, accommodation category (i.e., Presentation, Response, Timing/Scheduling, Setting, Check, Cues/Prompts, Structured Behavioral Interventions) across standardized measures of academic and behavioral functioning. The accommodation type variable was dichotomous, indicating whether a student did or did not receive at least one accommodation within each specific category (i.e., 0 = student received zero accommodations in the category, 1 = student received at least one accommodation within the category). The dependent variables were scores from the Parent BASC-2 (Externalizing subscale), MASC (total score), RADS-2 (total score), and WJ-III (Broad Reading and Broad Math clusters). Effect sizes were examined using Cohen's d. Again, accommodations used in the classroom and those designated for state/district-wide assessments were analyzed separately.

Results

As noted above, 1,840 accommodations were coded, 1,155 of which were applied in the classroom and 685 applied to state/district-wide assessments. Table 2 describes the 1,840 accommodations that were coded, noting the number and

Table 2. Frequency and Percentage of Students Who Received Each Accommodation Type.

Category and specific accommodation	Classroom (*n* = 222)	State/district assessment (*n* = 222)
Presentation	123 (55%)	123 (55%)
Read aloud (tests)	92 (41%)	115 (52%)
Copy of notes provided	40 (18%)	NA
Clear directions	35 (16%)	13 (6%)
Study guides	18 (8%)	1 (<1%)
Read aloud (classroom assignments)	16 (7%)	NA
Read aloud (classroom, not specified)	10 (5%)	NA
Read aloud (directions or instructions)	8 (4%)	25 (11%)
Response	98 (44%)	61 (27%)
Calculator	78 (35%)	49 (22%)
Graphic organizers	12 (5%)	1 (<1%)
Timing/Scheduling	182 (82%)	129 (58%)
Extended time (classroom, not specified)	41 (19%)	NA
Extended time (tests)	106 (48%)	125 (56%)
Breaks	25 (11%)	46 (21%)
Extended time (classroom assignments)	114 (51%)	NA
Chunking	31 (14%)	0 (0%)
Setting	177 (80%)	181 (82%)
Small group (testing)	106 (48%)	199 (90%)
Preferential seating	92 (41%)	28 (13%)
Small group (classroom not specified)	24 (11%)	NA
Small group (instruction)	22 (10%)	NA
Check	49 (22%)	6 (3%)
For understanding	22 (10%)	2 (1%)
Cues/Prompts	61 (27%)	26 (12%)
Cues (on-task behavior)	44 (20%)	13 (6%)
Cues (not specified)	17 (8%)	13 (6%)
Structured Behavioral Strategies	33 (15%)	0 (0%)
Positive reinforcement	21 (10%)	0 (0%)
Positive behavior support plan	15 (7%)	0 (0%)

percentage of students who received accommodations in each broad category, along with the specific type of accommodation (as noted verbatim on the spreadsheet), and whether the accommodation was applied in the classroom or on state/district-wide assessments. Note that most students received more than one accommodation (category and specific), so percentages exceed 100. Specific accommodations are displayed that were received by 5% or more students in either the classroom or on state/district-wide assessments. Among the 222 participants, the most frequent accommodations were Setting, provided to 176 (80%) students in the classroom and 180 (82%) students on state/district-wide assessments; Timing/Scheduling, provided to 182 (82%) students in the classroom and to 129 (58%) students on state/district-wide assessments; and Presentation, provided to 123 (55%) students in both the classroom and on state/district-wide assessments. All other accommodations were provided to fewer than 50% of students. The most frequent specific accommodation was Small Group Testing on state/district-wide assessments, provided to 90% of students, followed by Extended Time on state/district-wide assessments (provided to 56% of students), Read Aloud on state/district-wide assessments (provided to 52% of students), Extended Time on classroom tests (provided to 48% of students), and Small Group Testing on classroom tests (provided to 48% of students).

When comparing accommodation categories provided in the classroom with accommodations provided on state/district-wide assessments (see Table 2), Presentation and Setting accommodations were provided at equivalent or nearly equivalent rates. All other accommodations, however, were provided far less frequently on state/district-wide assessments than in the classroom. Specifically, Timing/Scheduling was received by 82% of participants in the classroom, but only 58% of participants on state/district-wide assessments. Similarly, Response was received by

Table 3. Frequency and Percentage of Students Within Disability Groups Who Received Accommodations.

Accommodation	Total sample (*N* = 221)	Disability category			
		SLD (*n* = 112)	EBD (*n* = 56)	OHI (*n* = 46)	Other (*n* = 7)
Presentation					
Classroom	122 (55%)	63 (56%)	33 (59%)	22 (48%)	4 (57%)
State/district assessment	122 (55%)	82 (73%)	16 (29%)	9 (20%)	5 (71%)
Response					
Classroom	98 (44%)	64 (57%)	17 (30%)	12 (26%)	5 (71%)
State/district assessment	61 (28%)	36 (32%)	10 (18%)	14 (30%)	1 (14%)
Timing/scheduling					
Classroom	181 (82%)	95 (85%)	46 (82%)	35 (76%)	5 (71%)
State/district assessment	128 (58%)	66 (59%)	33 (59%)	25 (54%)	4 (57%)
Setting					
Classroom	176 (80%)	87 (78%)	47 (84%)	36 (78%)	6 (86%)
State/district assessment	180 (81%)	97 (87%)	47 (84%)	31 (67%)	5 (71%)
Check					
Classroom	49 (22%)	25 (22%)	11 (20%)	13 (28%)	NA
State/district assessment	6 (3%)	3 (3%)	2 (4%)	1 (2%)	NA
Cues/prompts					
Classroom	61 (28%)	25 (22%)	22 (39%)	12 (26%)	2 (29%)
State/district assessment	26 (12%)	9 (8%)	13 (23%)	4 (9%)	NA
Structured behavioral strategies					
Classroom	33 (15%)	14 (13%)	12 (21%)	7 (15%)	NA
State/district assessment	NA	NA	NA	NA	NA

Note. Special education label information was missing for one student (*N* = 221). SLD = specific learning disability; EBD = emotional and behavioral disorder; OHI = other health impairment.

44% in the classroom and 27% of participants on state/district-wide assessments. Examining specific accommodations, a greater percentage of students received Tests Read Aloud, Extended Time (tests), Breaks and Small Group Testing, on state/district-wide assessments than within the classroom. Conversely, the percentages of students who received Clear Directions, Use of a Calculator, Graphic Organizers, Chunking, Preferential Seating, Check for Understanding, Cues (On-Task Behavior and Not Specified), Positive Reinforcement, and Positive Behavior Support Plans were greater within the classroom than on state/district-wide assessments.

The percentage of accommodations received in the classroom and on state/district-wide assessments by disability category is shown in Table 3. Overall, students in all disability categories received Timing/Scheduling and Setting accommodations in the classroom far more frequently than other types of accommodations (approximately 80% of students).

When comparing each disability category across accommodation type, similar percentages of students with SLD and EBD received accommodations in classroom for the areas of Presentation (56%, 59%, respectively) and Check (22%, 20%, respectively). Furthermore, similar percentages of students with SLD, EBD, and OHI received Timing (85%, 82%, 78%, respectively) and Setting (78%, 84%, 78%, respectively) classroom accommodations. Substantially greater differences were seen across disability categories for the provision of Response, Cues/Prompts, and Structured Behavioral Strategies as classroom accommodations.

On state/district-wide assessments, the percentage of students with SLD and EBD who received Timing/Scheduling was identical (59%) and was similar for Setting (87%, 84%, respectively), and (although far less frequent) Check (3%, 4%, respectively). The percentage of students with OHI was similar only for the accommodation of Timing/Scheduling (54%). All other accommodations on state/district-wide assessments varied considerably, depending on disability type.

The second research question examined differences in demographic characteristics between groups of students who did or did not receive at least one of each type of accommodation. With respect to classroom accommodations, chi-square tests for independence indicated a significant relationship between state of residence and Presentation accommodations, $\chi^2(4, N = 206) = 25.77, p < .001$, Cramer's $V = .354$; Setting accommodations, $\chi^2(4, N = 206) = 14.57, p = .006$, Cramer's $V = .266$; Structured Behavioral Intervention accommodations, $\chi^2(4, N = 206) = 33.66, p < .001$, Cramer's

$V = .404$; and Check accommodations, $\chi^2(4, N = 206) = 17.95$, $p = .001$, Cramer's $V = .295$. Adjusted residuals of ±2.0 were examined to determine cases where values were significantly different than expected. A significantly greater number of students in Missouri received Presentation accommodations than expected ($z_{adj} = 2.6$), whereas significantly fewer students in Kansas received Presentation accommodations ($z_{adj} = -4.4$). However, significantly fewer students than expected in Missouri received Setting accommodation in the classroom ($z_{adj} = -3.2$). In addition, students in Pennsylvania appear to be significantly more likely to receive a Structured Behavioral Intervention accommodation ($z_{adj} = 5.6$), especially compared with students in Ohio ($z_{adj} = -2.2$) and Kansas ($z_{adj} = -2.2$). Students in Pennsylvania also received Check accommodations in the classroom more frequently than expected ($z_{adj} = 3.4$), whereas students in Ohio received them less frequently than expected ($z_{adj} = -3.4$).

Separate analyses were conducted for accommodations on state/district-wide assessments. These chi-square results indicated significant associations between state of residence and Timing/Scheduling accommodations, $\chi^2(4, N = 189) = 138.82$, $p < .001$, Cramer's $V = .857$; Check accommodations, $\chi^2(4, N = 189) = 13.05$, $p = .011$, Cramer's $V = .263$; Response accommodations, $\chi^2(4, N = 189) = 20.58$, $p < .001$, Cramer's $V = .330$; and Cue/Prompt accommodations, $\chi^2(4, N = 189) = 29.05$, $p < .001$, Cramer's $V = .392$. Post hoc analysis of adjusted residuals indicated that students in Ohio received Timing/Scheduling accommodations on state/district-wide assessments far more frequently than expected ($z_{adj} = 6.6$), especially compared with students in South Carolina who received them significantly less frequently than expected ($z_{adj} = -11.6$). Although very few students across all states received any type of Check accommodation on state/district-wide assessments, adjusted residuals indicated that students in Pennsylvania received them significantly more frequently than expected ($z_{adj} = 2.7$). In addition, students in Kansas received Response accommodations more frequently than expected ($z_{adj} = 2.5$), whereas students in South Carolina received them less frequently than expected ($z_{adj} = -4.0$). Finally, adjusted residuals indicated that students in Pennsylvania received Cue/Prompt accommodations for state/district-wide assessments more frequently than expected ($z_{adj} = 4.7$), and students in South Carolina received them less frequently than expected ($z_{adj} = -3.3$).

When examining special education labels and classroom accommodations, chi-square results indicated a significant association between special education label type and Response accommodations, $\chi^2(3, N = 205) = 20.16$, $p < .001$, Cramer's $V = .314$. Specifically, students with an SLD label received Response accommodations in the classroom (e.g., calculator) more frequently than expected ($z_{adj} = 3.7$) compared with students with OHI ($z_{adj} = -2.7$) or EBD labels ($z_{adj} = -2.5$).

For accommodations provided on standardized tests, significant associations also were found between special education label type and Presentation accommodations, $\chi^2(3, N = 188) = 35.84$, $p < .001$, Cramer's $V = .437$; and Cue/Prompt accommodations, $\chi^2(3, N = 188) = 10.57$, $p = .014$, Cramer's $V = .237$. Students with an SLD label received Presentation accommodations on state/district-wide assessments significantly more frequently than expected ($z_{adj} = 5.0$), whereas students with an EBD label received this type of accommodation significantly less frequently than expected ($z_{adj} = -5.1$). In contrast, however, students with an EBD label received Cue/Prompt accommodations on state/district-wide assessments more frequently than expected ($z_{adj} = 3.2$), whereas students with an SLD label received this type of accommodation less frequently than expected ($z_{adj} = -2.1$).

Finally, chi-square tests indicated significant associations between student grade level and Presentation accommodations, $\chi^2(4, N = 189) = 10.03$, $p = .040$, Cramer's $V = .230$; and Check accommodations, $\chi^2(4, N = 189) = 11.23$, $p = .024$, Cramer's $V = .244$. Students in eighth and ninth grades received Check accommodations on state/local assessments more frequently than expected ($z_{adj} = 2.2, z_{adj} = 2.3$). However, students in 10th grade were more likely than expected to receive Presentation accommodations on state/district-wide tests, especially compared with students in ninth grade ($z_{adj} = -2.7$).

Cramer's V takes into account degrees of freedom. Therefore, slightly different criteria are recommended for judging the size of the effect for tables larger than 2 by 2. Cohen (1988) suggested that for chi-square analyses with 4 degrees of freedom, .25 is considered a large effect, .15 is considered a medium effect, and .05 is considered a small effect. For analyses with 3 degrees of freedom, .29 is considered a strong association, .17 is considered moderate, and .06 is considered a weak or negligible association. Following these criteria, all significant relationships appear to be moderate to strong (Cramer's $V = .23–.86$).

The third research question compared the number of accommodations that students received across demographics. In terms of accommodations received in the classroom, nonparametric analyses indicated that the only significant difference found among the demographic characteristics was related to state of residence. A Kruskal–Wallis Test revealed a statistically significant difference in the number of accommodations students received across the five states, $H(4) = 29.44$, $p < .001$. Pairwise comparisons with Bonferroni-adjusted p values showed significant differences between the number of classroom accommodations received by students in Pennsylvania ($Mdn = 6$, $n = 51$) compared with students in Ohio ($Mdn = 5$, $n = 50$, $p = .03$,

r = .30), Missouri (Mdn = 4.5, n = 18, p = .02, r = .37), South Carolina (Mdn = 4, n = 64, p < .001, r = .46), and Kansas (Mdn = 4, n = 23, p = .001, r = .45). No other significant differences were found between states. Kruskal–Wallis Tests revealed no significant differences related to primary disability category, $H(3) = 4.83$, p = .185; or grade level, $H(4) = 8.67$, p = .07. Furthermore, Mann–Whitney U Tests revealed no significant differences according to gender, U = 4,768, z = 1.13, p = .258, r = .08; or race, U = 4,945, z = −0.78, p = .436, r = .05.

Results from separate nonparametric analyses of state/district-wide assessment accommodations revealed nearly identical results, with a Kruskal–Wallis Test revealing a statistically significant difference in the number of accommodations students received depending on state of residence, $H(4) = 83.68$, $p < .001$. Pairwise comparisons with Bonferroni-adjusted p values showed significant differences between the number of standardized testing accommodations received by students in South Carolina (Mdn = 2, n = 51) compared with students in Ohio (Mdn = 4, n = 62, p < .001, r = .73), Missouri (Mdn = 3, n =15, p < .001, r = .51), Kansas (Mdn = 4, n = 19, p < .001, r = .72), and Pennsylvania (Mdn = 4, n = 42, p < .001, r = .78). No other significant differences were found between states. Again, Kruskall–Wallis Tests revealed no significant differences related to primary disability category, $H(3) = 2.16$, p = .54; or grade level, $H(4) = 4.26$, p = .371. Finally, Mann–Whitney U Tests revealed no significant differences according to gender, U = 3,193, z = −0.45, p = .652, r = .03; or race, U = 4,072.5, z = −1.02, p = .306, r = .07

The fourth research question compared academic and behavioral functioning between groups of students who did or did not receive at least one accommodation in each category. A series of independent samples t test of classroom accommodations revealed significant differences only among students who had Response accommodations, $t(170) = 2.87$, p = .005. Specifically, students who received a Response accommodation (e.g., calculator) in the classroom had significantly lower WJ-III Broad Math scores (M = 73.99, SD = 11.07) than students who did not receive a Response accommodation (M = 78.90, SD = 11.38). The magnitude of the differences in the means (mean difference = 4.91) demonstrated a medium effect (Cohen's d = 0.44). No significant differences were found on t-test analyses of other accommodation types based on academic or behavioral functioning.

Again, separate t tests were conducted for accommodations provided to students on state/district-wide assessments. Significant differences were found between groups of students who did or did not receive Presentation, Response, and Check accommodations. Students who received at least one Presentation accommodation on state/district-wide assessments (e.g., read aloud) had significantly lower WJ-III Broad Math scores (M = 73.25, SD = 10.22) and WJ-III Broad Reading scores (M = 80.39, SD = 11.16) than students who did not, WJ-III math, M = 79.48, SD = 11.21, $t(153) = 3.44$, p = .001; WJ-III reading, M = 90.49, SD = 10.79, $t(147) = 5.39$, p < .001. The magnitude of the differences in the means for WJ-III math scores (mean difference = 6.23) demonstrated a medium effect (Cohen's d = 0.59), whereas the means for WJ-III reading scores (mean difference = 10.10) demonstrated a very large effect (Cohen's d = 0.92). Similarly, students who received a Response accommodation on state/district-wide assessments had significantly lower WJ-III Broad Math scores (M = 70.94, SD = 10.12) than students who did not receive a Response accommodation, M = 77.20, SD = 10.74, $t(153) = 3.41$, p = .001. The magnitude of the differences in the means, mean difference = 6.27, 95% confidence interval (CI) = [2.64, 9.89], demonstrated a medium effect (Cohen's d = 0.59). Finally, students who received at least one Check accommodation on state/district-wide assessments demonstrated significantly greater risk of anxiety, as measured by the MASC (M = 58.17, SD = 7.57), than students who did not receive a Check accommodation, M = 48.99, SD = 10.49, $t(187) = -2.12$, p = .035. The magnitude of differences in means (mean difference = 9.178) demonstrated a large effect (Cohen's d = 0.88). Despite the statistically significant difference and large effect size, it is important to note that the mean total MASC score of students who received Check accommodations (M = 58.17) fell into the high average range and is just below the threshold for what is typically considered at risk of internalizing problems. On this measure, T-scores of 60 to 64 fall within the slightly elevated range, whereas scores of 65 to 69 fall within the elevated score range.

Discussion

To date, no empirical research studies have examined accommodations provided for high school students with disabilities either in the classroom or on standardized tests. This is a disturbing gap given the emphasis placed on standardized tests as well as the potential for accommodations to allow students to overcome performance obstacles resulting from their disability. The current study begins to fill this gap by shedding light on the types of accommodations high school students receive, the consistency between accommodations provided in the classroom and on state/district-wide assessments, and variables related to receipt of accommodations.

With respect to types of accommodations high school students with emotional and behavioral problems receive, in addition to the four categories previously described in the literature (i.e., Presentation, Response, Timing/Scheduling, Setting), three additional categories emerged that occurred relatively frequently across IEPs. Furthermore, frequent specific accommodations within these categories were listed on

IEPs and a large number of additional accommodations occurred infrequently and others were vague and could not be interpreted. These data underscore concerns raised by researchers (e.g., Harrison et al., 2013) regarding the immature status of research in this area and explain the challenges for teachers when attempting to identify appropriate accommodations for students (Fuchs & Fuchs, 2001).

When comparing accommodations provided in the classroom with accommodations provided on state/district-wide assessments, Presentation and Setting accommodations were provided at equivalent or nearly equivalent rates. All other accommodations, however, were provided far less frequently (or not at all) on state/district-wide assessments than in the classroom. For use of a calculator, it makes sense that this Response accommodation is not applied to state testing because the rules of the exams indicate when they may or may not be utilized. However, for others this is a highly concerning finding, in that, conceptually similar benefit should be accrued from a particular accommodation, regardless of how it is applied. In fact, numerous sources recommend that accommodations be similarly applied. For example, The Pennsylvania Accommodations Guidelines (2018) state, "The student must be provided the selected accommodations during instructional periods that necessitate their use. An accommodation may not be used solely during assessments" (p. 21). Increased attention to classroom accommodations is particularly important for students with emotional and behavioral problems due to long documented challenges in school performance (Newman et al., 2011; Simpson et al., 2011).

The data revealed variability in some types of accommodations relative to disability label. Overall, similar percentages of students with SLD and EBD received each type of accommodation. These findings are inconsistent with previous literature indicating teachers are more reluctant to provide accommodations to students with EBD than students with other disability labels (Maccini & Gagnon, 2006). However, Maccini and Gagnon reported results of a survey, rather than directly comparing disability types. Another explanation is that the students in our survey with SLD also engaged in behavior problems. Findings might align with Maccini and Gagnon when comparing a sample of students with SLD who do not exhibit behavior problems.

Perhaps the most notable finding in this analysis is that the percentage of students with OHI who received accommodations in both the classroom and on state/district-wide assessments was smaller across almost every accommodation type compared with students with SLD and EBD labels. This is difficult to explain because inclusion in the larger CARS study required that students exhibit significant emotional and/or behavioral problems as well as school impairment, suggesting that all participants would have benefited from accommodations. These findings support the need for research to examine why students with OHI receive so few accommodations relative to other disability groups and what accommodations would be most beneficial.

Our second and third research questions examined the relationship between demographic variables and type and number of accommodations. Several categories as well as the number of accommodations were significantly different depending on students' state of residence. One possible explanation is that schools and/or districts tend to assign similar accommodations, which elevated similarities within states. Another explanation might be differences in state regulations. For example, The Pennsylvania System of School Assessment (PSSA), required throughout Pennsylvania, is not a timed test and students may be provided extra time (beyond that scheduled) to complete the assessment, provided it is adjacent to the testing period (2011–2012 PSSA Assessment Handbook). This stipulation could reduce the number of Timing/Scheduling accommodations on the IEPs of students in Pennsylvania. This might also explain the lower percentage of students receiving Timing/Scheduling accommodations on state/district-wide assessments compared with the classroom.

Differentiated accommodation type also was related to special education label. In the absence of in-depth analyses of specific disability symptomatology and the relationship to accommodation effectiveness, it is difficult to ascertain whether this is appropriate. Similarly, the rationale for grade-level differences is difficult to interpret. It may be that teachers become more reluctant to provide certain accommodations, such as checking for understanding, as students age due to expectations for increasing independence. However, this does not explain differences in Presentation accommodations. No differences were found for gender or race.

It is also unclear why the number of accommodations significantly differed by state of residence. Although not supported by research, it is possible that states are discrepant in their emphasis on accommodation provision, which also includes the belief that a greater number of accommodations is likely to address obstacles to performance.

The final research question pertained to differences in accommodations relative to academic, emotional, or behavioral functioning as measured by standardized tests. We found few significant differences. The only difference found for classroom accommodations was students with a lower Broad Math score on the WJ-III were significantly more likely to receive a Response accommodation (the most frequent being use of a calculator). On state/district-wide assessments, students with lower Broad Math and Broad Reading scores WJ-III were significantly more likely to receive a Presentation accommodation, students with lower Broad Math scores WJ-III were more likely to receive a Response accommodation, and students with higher anxiety

scores on the MASC were more likely to receive a Check accommodation. The relationship between Response accommodations and academic functioning is logical. More broadly, the fact that so few accommodations were linked to these three areas of functioning was unexpected. That is, it seems that performance on standardized tests should be directly related to particular types of accommodations. This is the case for accommodations that pertain more directly to academic performance (e.g., study guide, read aloud) and those that focus primarily on behavior (Cues/Prompts, positive reinforcement). Furthermore, the only accommodation provided significantly more frequently for students with anxiety was Check on state/district-wide assessments. The most common form of Check was for understanding. It is questionable whether there is evidence that checking for understanding during standardized testing is beneficial for students with anxiety. In general, these data suggest that students may be receiving accommodations that are not well matched to student needs.

Although the current study initiates a literature base in the area of accommodations for secondary school students with disabilities, several limitations warrant discussion. First, the sample was drawn from a larger study in which students needed to meet inclusion criteria. Thus, the sample may not be representative of all students with disabilities who have emotional and behavioral problems. Nonetheless, due to our recruitment and screening procedure, our sample captured students who school staff viewed as having the most significant problems. Therefore, our data may be more reflective of accommodations selected for students with disabilities who engage in behavior problems than those who no longer exhibit problems but have retained the label. Furthermore, although the overall sample size was robust, there was a relatively small sample of students within some disability categories (e.g., OHI [$n = 46$]). Caution should be used when examining descriptive data (see Table 1).

In addition, as noted above, some states allow changes (e.g., extended time) on state/district-wide assessments that would typically be considered an accommodation. Thus, accommodations listed on IEPs may be an underestimate of the actual number in some states. In addition, state assessments may not be administered every year. Again, accommodations that students typically receive on state/district-wide assessments may not be reflected during years without testing. This, however, should not have affected classroom accommodations.

Another limitation pertains to the definition and classification of an accommodation. We coded accommodations listed on IEPs without judging whether they should meet the technical definition of an accommodation. For example, Structured Behavioral Strategies, defined as systematic behavioral intervention approaches (e.g., positive behavior support plans, consistent use of positive reinforcement for following expectations) appeared fairly frequently on IEPs as an accommodation. However, this might be more appropriately considered part of a behavior support plan. Behavior support plans are sometimes written and stored independent of the IEP and it is possible that they are less accessible to general education teachers. Although speculative, it may be that special education teachers include behavior supports in the IEP to increase the likelihood they will be implemented by general education teachers.

Findings from the current study suggest that overall accommodation selection is poorly understood and, similar to previous research (Fuchs & Fuchs, 2001), a highly imprecise practice. Furthermore, the multitude of accommodations that students received (range = 1–27) suggest a lack of clarity about effective accommodations and perhaps the need for multiple accommodations to address a student's disability. Issues with accommodation selection are further supported by the mismatch between student performance on standardized tests and specific accommodations. Additional research is needed regarding accommodation selection and whether multiple accommodations offer an additive benefit. Also, the large number that fell into the Other category and the new categories that emerged suggest the need to further explore the manner in which accommodations are developed for students with emotional and behavioral problems as well as their effectiveness.

Authors' Note

The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Institute of Education Sciences, U.S. Department of Education, Grant R324C080006, awarded to Lehigh University.

References

Accommodations Guidelines. (2018). Commonwealth of Pennsylvania, Department of Education. Retrieved from http://www.education.pa.gov/Documents/K-12/Assessment%20and%20Accountability/PSSA/Accommodations/2018%20Accommodations%20Guidelines%20Handbook.pdf

Agresti, A. (2013). *Categorical data analysis* (3rd ed.). Hoboken, NJ: John Wiley.

Bradley, R., Doolittle, J., & Bartolotta, R. (2008). Building on the data and adding to the discussion: The experiences and outcomes of students with emotional disturbance. *Journal of Behavioral Education*, *17*, 4–23. doi:10.1007/s10864-007-9058-6

Carr-George, C., Vannest, K. J., Willson, V., & Davis, J. L. (2009). The participation and performance of students with emotional and behavioral disorders in a state accountability assessment in reading. *Behavioral Disorders*, *35*, 66–78.

Christensen, L. L., Braam, M., Scullin, S., & Thurlow, M. L. (2011). *2009 state policies on assessment participation and accommodations for students with disabilities* (Synthesis Report 83). Minneapolis: National Center on Educational Outcomes, University of Minnesota.

Christenson, S. L., Decker, D. M., Triezenberg, H. L., Ysseldyke, J. E., & Reschly, A. (2007). Consequences of high-stakes assessment for students with and without disabilities. *Educational Policy*, *21*, 662–690.

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cramer, H. (1999). *Mathematical methods of statistics*. Princeton, NJ: Princeton University Press.

Elliott, S. N., & Marquart, A. M. (2004). Extended time as a testing accommodation: Its effects and perceived consequences. *Exceptional Children*, *70*, 349–367. doi:10.1177/001440290407000306

Feldman, E., Kim, J. S., & Elliott, S. N. (2011). The effects of accommodations on adolescents' self-efficacy and test performance. *The Journal of Special Education*, *45*, 77–88. doi:10.1177/0022466909353791

Field, A. (2013). *Discovering statistics using SPSS* (4th ed.). London, England: SAGE.

Fuchs, L. S., & Fuchs, D. (2001). Helping teachers formulate sound test accommodation decisions for students with learning disabilities. *Learning Disabilities Research & Practice*, *16*, 174–181. doi:10.1111/0938-8982.00018

Fuchs, L. S., Fuchs, D., & Capizzi, A. M. (2005). Identifying appropriate test accommodations for students with learning disabilities. *Focus on Exceptional Children*, *37*(6), 1–8.

Fuchs, L. S., Fuchs, D., Eaton, S. B., Hamlett, C., Binkley, E., & Crouch, R. (2000). Using objective data sources to enhance teacher judgments about test accommodations. *Exceptional Children*, *67*, 67–81. doi:10.1177/001440290006700105

Fuchs, L. S., Fuchs, D., Eaton, S. B., Hamlett, C., & Karns, K. M. (2000). Supplementing teacher judgments of mathematics test accommodations with objective data sources. *School Psychology Review*, *29*, 65–85.

Gage, N. A., Wilson, J., & MacSuga-Gage, A. S. (2014). Writing performance of students with emotional and/or behavioral disabilities. *Behavioral Disorders*, *40*, 3–14.

Harrison, J. R., Bunford, N., Evans, S. W., & Owens, J. S. (2013). Educational accommodations for students with behavioral challenges: A systematic review of the literature. *Review of Educational Research*, *83*, 551–597. doi:10.3102/0034654313497517

Helwig, R., & Tindal, G. (2003). An experimental analysis of accommodation decisions on large-scale mathematics tests. *Exceptional Children*, *69*, 211–225. doi:10.1177/001440290306900206

Individuals With Disabilities Education Act (IDEA), 20 U.S.C. §§ 1400–1419 (2004).

Kern, L., Evans, S. W., Lewis, T. J., State, M., Weist, M. D., & Wills, H. P. (2015). CARS comprehensive intervention for secondary students with emotional and behavioral problems: Conceptualization and development. *Journal of Emotional and Behavioral Disorders*, *23*, 195–205. doi:10.1177/1063426615578173

Ketterlin-Geller, L. R., Yovanoff, P., & Tindal, G. (2007). Developing a new paradigm for conducting research on accommodations in mathematics testing. *Exceptional Children*, *73*, 331–347. doi:10.1177/001440290707300304

Lang, S. C., Elliott, S. N., Bolt, D. M., & Kratochwill, T. R. (2008). The effects of testing accommodations on students' performances and reactions to testing. *School Psychology Quarterly*, *23*, 107–124. doi:10.1037/1045-3830.23.1.107

Maccini, P., & Gagnon, J. C. (2006). Mathematics instructional practices and assessment accommodations by secondary special and general educators. *Exceptional Children*, *72*, 217–234. doi:10.1177/001440290607200206

March, J. (1998). *Manual for the Multidimensional Anxiety Scale for Children (MASC)*. Toronto, Ontario, Canada: Multi-Health Systems.

March, J. S., Sullivan, K., & Parker, J. (1999). Test-retest reliability of the Multidimensional Anxiety Scale for Children. *Journal of Anxiety Disorders*, *13*, 349–358.

McKevitt, B. C., & Elliott, S. N. (2003). Effects and perceived consequences of using read-aloud and teacher-recommended testing accommodations on a reading achievement test. *School Psychology Review*, *32*, 583–601.

Newman, L., Wagner, M., Knokey, A. M., Marder, C., Nagle, K., Shaver, D., & Wei, X. (2011). *The post-high school outcomes of young adults with disabilities up to 8 years after high school: A report from the National Longitudinal Study-2 (NLTS2)*. Menlo Park, CA. Retrieved from http://www.nlts2.org/reports/2011_09_02/nlts2_report_2011_09_02_complete.pdf

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children* (2nd ed.). Circle Pines, MN: American Guidance Service.

Reynolds, W. M. (2002). *Reynolds Adolescent Depression Scale: Professional manual* (2nd ed.). Lutz, FL: Psychological Assessment Resources.

Schulte, A. A. G., Elliott, S. N., & Kratochwill, T. R. (2001). Effects of testing accommodations on standardized mathematics test scores: An experimental analysis of the performances of students with and without disabilities. *School Psychology Review*, *30*, 527–547.

Simpson, R., Peterson, R., & Smith, C. (2011). Critical educational components for students with emotional and behavioral disorders: Science, policy, and practice. *Remedial and Special Education*, *32*, 230–242. doi:10.1177/0741932510361269

Temple-Harvey, K. K., & Vannest, K. J. (2012). Participation and performance of students with emotional disturbance on a statewide accountability assessment in math. *Remedial and Special Education*, *33*, 226–236. doi:10.1177/0741932510391815

Thaler, N. S., Kazemi, E., & Wood, J. J. (2010). Measuring anxiety in youth with learning disabilities: Reliability and

validity of the Multidimensional Anxiety Scale for Children (MASC). *Child Psychiatry & Human Development*, *41*, 501–514.

U.S. Department of Education, National Center for Education Statistics. (2016). *Digest of Education Statistics: 2015* (NCES 2016-014). Retrieved from https://nces.ed.gov/programs/digest/d15/ch_2.asp

Wagner, M., Friend, M., Bursuck, W. D., Kutash, K., Duchnowski, A. J., Sumi, W. C., & Epstein, M. H. (2006). Educating students with emotional disturbances: A national perspective on school programs and services. *Journal of Emotional and Behavioral Disorders*, *14*, 12–30. doi:10.1177/10634266060140010201

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.

Ysseldyke, J., Nelson, J. R., Christenson, S., Johnson, D. R., Dennison, A., Triezenberg, H., . . . Hawes, M. (2004). What we know and need to know about the consequences of high-stakes testing for students with disabilities. *Exceptional Children*, *71*, 75–95. doi:10.1177/001440290407100105

Ysseldyke, J., Thurlow, M., Bielinski, J., House, A., Moody, M., & Haigh, J. (2001). The relationship between instructional and assessment accommodations in an inclusive state accountability system. *Journal of Learning Disabilities*, *34*, 212–220. doi:10.1177/002221940103400302