SAGAMORE JOURNALS

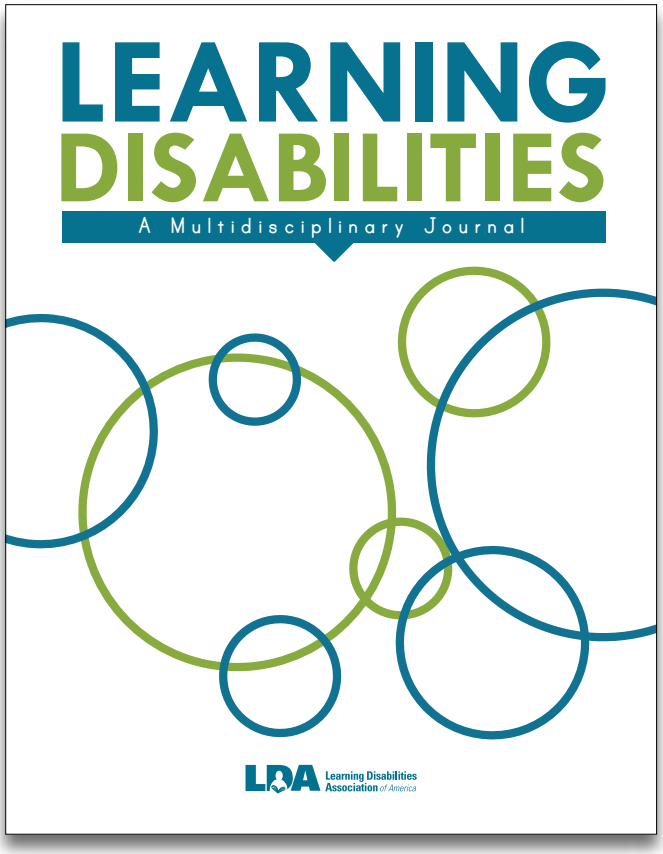
FREE ARTICLE PREVIEW

Learning Disabilities: A Multidisciplinary Journal

Learning Disabilities: A Multidisciplinary Journal (LDMJ) is an official publication of the Learning Disabilities Association of America (LDA). The journal is a vehicle for disseminating the most current thinking on learning disabilities and to provide information on research, practice, theory, issues, and trends regarding learning disabilities from the perspectives of varied disciplines involved in broadening the understanding of learning disabilities for both children and adults. LDA and ***LDMJ*** are committed to the study of learning disabilities through all stages of life. The disciplines represented in the journal include, but are not limited to, the following:

- advocacy, law, public policy;
- assessment, research, technology, science, and trends;
- early childhood education;
- higher and adult education;
- vocational and career education;
- cultural differences;
- public and private education;
- family dynamics; and
- mental health, social and emotional issues, and social work.

Since 1963, LDA has provided support to people with learning disabilities and their parents, teachers, and other professionals with cutting-edge information on learning disabilities, practical solutions, and a comprehensive network of resources. These services make LDA the leading resource for information on learning disabilities.



Screening to Assign Accommodations: Using Data to Make Decisions

Leanne R. Ketterlin-Geller, Lindy Crawford, and Jacqueline N. Huscroft-D'Angelo

In this paper, we describe the design, development, and initial validity evidence for the Screening to Assign Accommodations Tool (SAAT). The purpose of the SAAT is to identify students for whom construct irrelevant variance (CIV) might obscure measurement of their knowledge, skills, and abilities in mathematics. Once students are identified as having risk factors that might interfere with accurate measurement of the tested construct, additional information can be collected to determine the most appropriate accommodations that will reduce the impact of construct irrelevant variance on their observed academic performance. Implications, limitations, and future research for the refinement of this instrument are discussed.

Keywords: *Testing accommodations; accommodations assignment; mathematics assessment*

Test accommodations are changes to the presentation, equipment, or materials; response mode, setting, and timing; or scheduling of a test that are intended to mediate individual characteristics that might cause construct irrelevant variance (CIV) in students' scores. CIV is systematic error in students' test scores that is caused not by their knowledge, skills, and abilities in the tested domain but instead by their disabilities manifesting in ways that interfere with the student's ability to interact fully with the tested content (Haladyna & Downing, 2004).

Test accommodations support accurate measurement of students' knowledge, skills, and abilities in the tested construct. In a study conducted by Santoli, Sachs, Romey, and McClurg (2008), 98.2% of special educators were willing to provide accommodations for students with disabilities to support their access to the general education curriculum. Although educators appear positive about the use of accommodations, teachers' decisions to assign accommodations may lack accuracy and specificity as related to individual student's needs (Crawford & Ketterlin-Geller, 2013; Rogers, Christian, & Thurlow, 2012). This is particularly salient for students with learning disabilities (Elbaum, Arguelles, Campbell, & Saleh, 2004).

Need for Assessment of Accommodation Assignment

Students with learning disabilities often display slow processing speed, visual-motor deficits, or attention difficulties that may result in challenges with written expression, reading fluency, or mathematical calculations. These characteristics may inadvertently impact educators' judgment about the need for accommodations, potentially leading to inappropriate accommodations assignment (Elbaum et al., 2004). Inappropriate assignment occurs when a student's characteristics are not sufficiently supported to reduce the impact of CIV, and includes over assignment (i.e., an accommodation is not warranted given the student's characteristics) and under assignment (i.e., an accommodation is not provided that may mitigate the negative impact of the student's characteristics) (Crawford & Ketterlin-Geller, 2013). Teachers' lack of knowledge and training on providing accommodations is often cited as the cause of inappropriate assignment (Fuchs & Fuchs, 2001; Ketterlin-Geller, Alonzo, Braun-Monegan, & Tindal, 2007; Maccini & Gagnon, 2006).

Appropriate assignment of accommodations is requisite for mediating individual characteristics that can lead to CIV, and can directly impact how students with disabilities access the general education curriculum and perform on assessments. Inappropriate assignment of accommodations can be moderated by the use of performance data during accommodation assignment procedures (Edgemon, Jablonski, & Lloyd, 2006). Fuchs and Fuchs (2001) developed a system by which teachers administer alternate forms of short tests using different accommodations. For example, a student takes a mathematics concepts and applications test in the standard

format as well as with extended time, a calculator, and an adult reading the text aloud. The experimental data from these varying conditions is evaluated to determine if the student demonstrated a differential boost in his or her performance as a result of the accommodation. These researchers found that data about students' response to accommodations on short tests can be generalized to large-scale achievement tests. Fuchs, Fuchs, Eaton, Hamlett, Binkley, and Crouch (2000) emphasize the importance of using this type of individualized approach to identify accommodations for students with learning disabilities because of the heterogeneity of their needs and the interwoven nature of their disability and the tested constructs. Although the diagnostic information generated by this individualized approach may be useful for assigning accommodations, students are required to take multiple assessments often demanding significant instructional time and resources. We contend that screening for possible accommodations can lead to valid decisions about students' need for accommodations, and increase the cost effectiveness and time efficiency of this process.

The purpose of this paper is to propose a screening tool that can be used to identify students whose disabilities may manifest in ways that interfere with accurate measurement of their knowledge, skills, and abilities in mathematics. The Screening to Assign Accommodations Tool (SAAT) incorporates multiple data sources to determine the student's characteristics that can lead to CIV. Results from the SAAT include recommended accommodations that can reduce the impact of CIV. Once these accommodations are identified, an experimental approach should be implemented to verify and provide validity evidence for their use.

What is the Screening Process?

In general, screening is defined as a process that proactively identifies the relative needs of an individual based on a cost effective and time efficient assessment of specific risk factors (Glover & Albers, 2007). As opposed to waiting until the problem escalates to provide support, screening allows for the early detection of possible problems followed by targeted treatment to prevent further issues or minimize the impact of the problem on the individual (Albers, Glover, & Kratochwill, 2007). Many people are familiar with the process of screening as it relates the medical field. A common example is a "well-baby check-up" in which a trained medical professional evaluates common indicators that are predictive of growth and development in young children. An individual child's data are compared to normative information to determine if the child is growing and developing as expected. If there is a discrepancy between the observed and expected data, the medical professional can gather additional information to make diagnostic and treatment decisions. As this example indicates, screening is often viewed as a prerequisite for intervening with treatment.

Screening decisions are facilitated by the use of a screening instrument or screener. Because screening is used to identify individuals with risk factors for developing a problem, results from a screener should differentiate between individuals with varying levels of risk, often along a continuum (Albers et al., 2007). A screener should be cost effective, time efficient, provide accurate information with acceptable levels of technical adequacy, and incorporate multiple sources of data from various perspectives (Glover & Albers, 2007).

Screeners aid in making decisions for various problems in medicine, mental health, and education. In education, screeners help educators make timely decisions to support students who are at risk for academic or behavioral difficulties. In some cases, these decisions are made within a well-established multitiered system of support (i.e., response to intervention) in which the problem is well specified and appropriate solutions are available. However, the screening process can be applied outside of this framework to help educators make meaningful decisions. The purpose of this article is to document the development of the SAAT to support teachers' assignment of accommodations.

Framework for the SAAT Development

Because the SAAT is a novel approach to early detection of a problem (i.e., presence of factors that can lead to CIV), the process of test development can be applied to design, develop, and evaluate this screener. To aid in the development of the SAAT, we adapted the systematic 12-step approach to test development proposed by Downing (2006; see Figure 1). Due to differences between traditional test development efforts proposed by Downing and the development of screening instruments, we incorporate fewer steps in the following framework (Note: steps 6, 7, 9, 11, and 12 are not included):

Step 1: Construct Definition

Identify the adverse outcome. What adverse outcome will the screener measure? An adverse outcome (or problem) is identified in which intervention can be applied, and serves as the basis for interpreting the results of the screener.

Step 2: Content Definition

What content will be assessed to make interpretations about the adverse outcome? Observable factors are identified that lead to the adverse outcome, and the

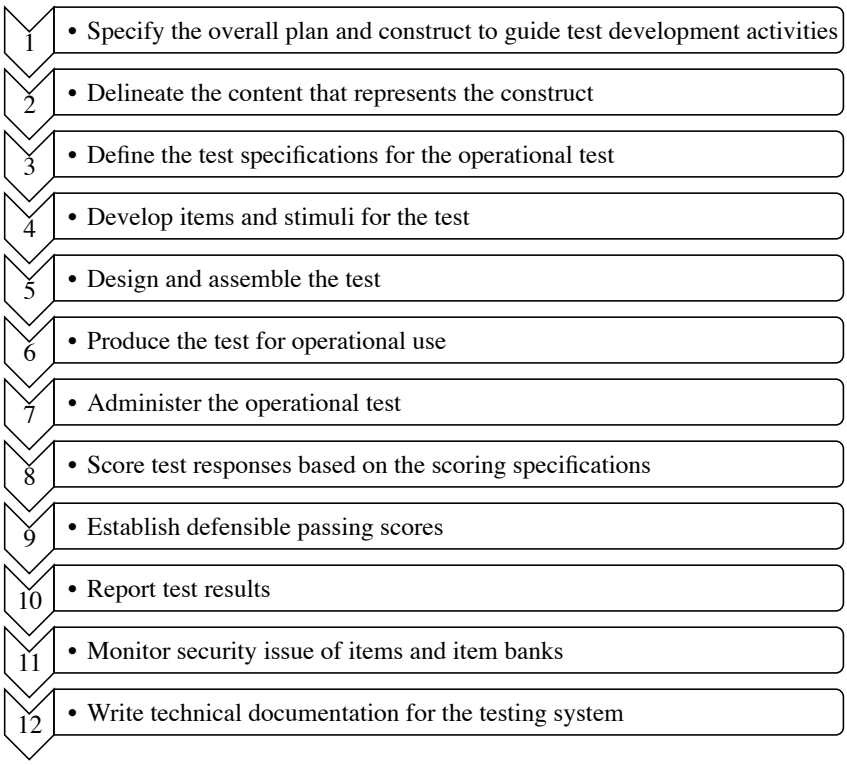


Figure 1. Twelve-step test development approach as specified by Downing (2006).

contribution of these observable factors is empirically examined to determine the relative importance of each factor to an individual's risk for developing the adverse outcome.

Steps 3-4: Test and Item Specifications

Define the test and item format, administration and scoring, and informants for assessing risk factors.

Details are specified for the screening tool to evaluate an individual's risk on the observable factors. Data are gathered from relevant individuals on the risk factors.

Step 5: Design and Assemble the Test

Format the operational test to maximize the collection of reliable data.

Steps 8, 10: Scoring Responses and Reporting Results

Assess an individual's level of risk. Score, report, and interpret the individual's level of risk for the adverse outcome.

To use the results of a screener within a prevention-intervention system, additional steps can be taken. Specifically, for individuals identified as being at risk for developing the adverse outcome, diagnostic information is gathered to determine the level of intervention needed. Using the results of the diagnostic assessment as a guide, appropriate interventions should be applied with the intent of changing the adverse outcome (Glover & Albers, 2007). An individual's response to the intervention may vary. Therefore, the individual's response to the intervention should be monitored over time.

In this article, we apply this process to identify individuals for whom CIV might obscure accurate measurement of knowledge, skills, and abilities. We propose the SAAT as a screener to identify student characteristics that may act as risk factors for CIV. We connect these risk factors to plausible test accommodations intended to mediate the impact of CIV on measurement of students' mathematics abilities. Finally, we situate this screening process within a systemic data gathering system that can be used to improve teachers' assignment of accommodations

Development of the SAAT

The SAAT identifies students for whom construct irrelevant variance (CIV) might obscure measurement of their knowledge, skills, and abilities in mathematics. The tool was developed over an extended time period using the framework described above. The application and evidence collected to justify the use of the SAAT for making accommodations assignment decisions is documented below.

Step 1: Construct Definition: CIV as an Adverse Outcome

When designing a test, the first step is to define the purpose of the assessment in relation to the construct to be measured (Downing, 2006). For a screener, "one of the most important implications of [screening] models is that a set of [risk] factors and the interaction amongst them often precede the manifestation of a full-blown problem" (Elliott, Huai, & Roach, 2007, p. 138). As such, the construct should align with theory and/or empirical data examining characteristics of risk that are associated with the problem of interest (Glover & Albers, 2007). Therefore, the purpose of the SAAT is to identify students who may be at risk for having CIV interfere with measurement of their mathematics knowledge, skills, and ability.

CIV is caused by irrelevant, but stable, personal characteristics that impact a student's ability to engage with the tested construct (Haladyna & Downing, 2004). These characteristics can lead to predicted patterns of interference in a student's ability to access the tested content, attend to the stimuli, or respond in a manner that allows the student to accurately demonstrate his or her knowledge. For example, a student with limited reading proficiency might not be able to decode the text used to present a mathematics item. The student's reading ability interferes with his or her ability to understand and respond to the item in a way that demonstrates his or her mathematics knowledge (or lack thereof). As this example illustrates, an adverse situation in which educators seek to intervene is inaccurate measurement due to CIV. The SAAT is intended to identify students' personal characteristics that interfere with accurate measurement.

Step 2: Content Definition: What Are the Risk Factors of CIV?

After the purpose has been articulated relevant to the construct, the next stage of test development is defining the content domain from which items will be sampled (Downing, 2006). The content of a screener should be based on the conceptualization of risk factors and account for multiple risk factors as well as factors that are associated with more than one problem (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007). Moreover, for screening measures to accurately discriminate between individuals who may and may not be at risk for the adverse outcome, the relative contribution of each factor to the individual's overall risk needs to be evaluated.

Students with disabilities present a variety of personal characteristics that might interfere with accurate measurement of their knowledge, skills, and abilities. Although additional research is needed to investigate risk factors that can contribute to CIV, some observable factors are known to interfere with accurate measurement including an individual's short and long-term memory and/or attention span, as well as linguistic and/or physical characteristics. It is important to note that students with disabilities have varying learner characteristics that may extend beyond these factors. Moreover, students with LD in mathematics may exhibit additional risk factors also relevant to the measured construct.

Although an empirical evaluation of the risk factors for CIV is not available, research on the use of accommodations leads to some plausible conclusions about the contribution of each factor. When interpreting accommodations research, evidence points to the possibility of CIV if a differential boost in performance is observed for students with disabilities when using the accommodation, as compared to students without disabilities under the same conditions. If improvements in performance are observed for both students with and without disabilities when using an accommodation, the accommodation may not be reducing causes of irrelevant variance (Fuchs & Fuchs, 2001).

We reviewed research on the effectiveness of accommodations for mediating the impact of students' personal characteristics on CIV to generate plausible conclusions about the contribution of risk factors. The research on accommodations research described below relies extensively but not exclusively on studies reviewed by the National Center on Educational Outcomes (NCEO) documented in three summaries of research dating between 2005-2010 (Cormier, Altman, Shyyan, & Thurlow, 2010; Rogers, Christian, & Thurlow, 2012; Zenisky & Sireci, 2007). These research summaries represent 120 publications meeting the high-quality standards set by NCEO.

Poor working or long-term memory. Students with learning disabilities often exhibit memory difficulties such as limited retention of information and/or difficulty accessing information stored in long-term memory (Geary, Hoard, Nugent, & Byrd-Craven, 2007). When interacting with items on a mathematics test, students with such difficulties may not be able to retrieve basic computation facts, may struggle to recall specific solution strategies, and may have difficulty distinguishing between relevant and irrelevant information. In some cases, these personal characteristics intentionally interact with the tested construct. For example, in tests of computational fluency, a student's ability to retrieve basic computational facts is the focus of measurement. However, when the intent of measurement is not related to these dimensions, these memory difficulties serve as risk factors for CIV.

To address poor working or long-term memory in mathematics tests, the most common accommodation is the use of a calculator. As an accommodation, calculators are intended to support students with memory or cognitive difficulties that impact their ability to recall computational facts or execute mathematical procedures. This accommodation is appropriate on tests in which the intent is to assess content other than a student's ability to perform operations. Research reviewed by Rogers et al. (2012) between 2009–2010 indicates that students with disabilities did not differentially benefit from using graphing and four-function calculators. Specifically, of the three studies included in their review, students with disabilities scored similarly as students without disabilities when using a calculator.

Graphic organizers have not been extensively studied as a test accommodation for students with learning disabilities, however, recent research indicates that providing graphic organizers along with a metacognitive structure for problem solving for students with mathematical difficulties during instruction significantly improves their learning of the content (c.f., Ives, 2007; Jitendra et al., 2009). Although not investigated as a testing accommodation, students with disabilities were better able to execute solution strategies as well as distinguish relevant from irrelevant information in the mathematics problem. Use of a graphic organizer as a test accommodation to support students with poor working or long-term memory should be empirically evaluated prior to implementation.

Attention difficulties. Studies indicate that anywhere from 40-80% of students with disabilities experience difficulty with sensory processing (Ahn, Miller, Milberger, & McIntosh, 2004; Talay-Ongan & Wood, 2000); these processing challenges likely make attending to mathematics items or tasks difficult. For example, some students demonstrate sensitivity to certain stimuli (e.g., ambient noise or fluorescent lights), or become over stimulated in certain contexts (e.g., loud and active classrooms). In some cases, presenting multiple items or tasks on one page may be over stimulating. Because the ability to attend to the test is peripheral to the tested construct, attention difficulties act as risk factors for CIV.

Testing in a separate location has been used as an accommodation for students who are easily distracted. However, of the three studies conducted between 2009–2010 that included individual administration, no study examined the effect of individual administration on students' scores in isolation of other accommodations (Rogers et al., 2012); therefore it is not possible to disentangle the effect of this accommodation from that of other changes.

Extended time is another accommodation often employed for students who are over-stimulated in a typical classroom. Although extended time was studied more than any other accommodation during 2005–2006 (seven out of 21 experimental or quasi-experimental studies summarized) and found to have positive effects for students with disabilities (Zenisky et al., 2007), both Rogers et al. (2012) and Cormier et al. (2010) reported a decrease in studies examining extended time. Moreover, mixed effects were observed, with some studies reporting a differential boost in students' scores with extended time, and other studies reporting no boost. Cormier et al. also notes that extended time is now incorporated into most state assessment procedures for all students, making its use as an accommodation less imperative. Looking forward, neither the Smarter Balanced Assessments (2012) nor the Partnership for Assessment of Readiness for College and Careers (PARCC; 2013) Assessments have included strict time demands in their administration guidelines. The Smarter Balanced Assessments are designed as untimed tests, and the PARCC Assessment provides students a set amount of additional time for each session as needed.

Linguistic or language difficulties. Students with learning disabilities may experience linguistic or language difficulties that include difficulty comprehending text or spoken language, understanding non-mathematical vocabulary, or expressing oneself (Butterworth & Reigosa as cited in Berch & Mazzocco, 2007; Moschkovich, 2010). If these symptoms are pronounced, the student may be receiving speech-language therapy or other support services. However, some students with learning disabilities demonstrate these difficulties but do not receive specialized therapy or services. When taking a mathematics test, these difficulties might interfere with a student's ability to understand the directions or prompts, know the meaning of the words contained in a word problem, or respond in the expected manner. Presence of linguistic or language difficulties might be risk factors for CIV.

Possible accommodations for students with linguistic or language difficulties include reading aloud mathematics test items and/or directions, using simplified language, and clarifying directions. A read aloud accommodation is often used for students who have poor decoding skills or other language challenges. As a presentation accommodation, the read aloud accommodation was studied more than any other accommodation between 2007–2010 (Cormier et al., 2010; Rogers et al., 2012). Read aloud accommodations may be provided by a human, a computer, or through a taped recording. Rogers et al. reported favorable findings for the read aloud accommodation with the majority of studies reporting a differential boost.

Accommodations focused on the linguistic features of test items have also been studied as a method for leveling the playing field for students with language difficulties. Although these accommodations have not been studied extensively, findings have not favored use of these accommodations for students with language difficulties. In summarizing results of a study published by Shaftel, Belton-Kocher, Glasnapp, and Poggio (2006), Zenisky et al. (2007) reports "Linguistic features of items have a greater effect for younger students, but no impact was found for students with disabilities" (p. 44). Similarly, in a study conducted by Ketterlin-Geller, Yovanoff, and Tindal (2007), as reported in Cormier et al. (2010), use of simplified language in the presentation of complex mathematics items with complex linguistic features showed no differential boost for students with reading difficulties.

Physical attributes. Students with disabilities may also have physical attributes that impact the way in which they interact with the environment including visual, hearing, orthopedic, neuromuscular, cardiovascular, and pulmonary disorders. The challenges caused by these attributes may impact a student's ability to gain physical access to the classroom, testing center, and/or content, and may result in decreased hand-eye coordination or limited fine or gross motor skills, may limit a person's ability to communicate verbally, and may reduce an individual's stamina or endurance. Due to federal legislation including the Americans with Disabilities Act of 1990, most public education buildings have been retrofitted to address issues that limited students' physical access to classrooms. However, depending on the individual student's needs, physical access to a classroom may not be accompanied by full access to the tested content. When taking a mathematics test, a student's physical limitations may interfere with accurate measurement, and therefore serve as risk factors for CIV.

Research on the effects of accommodations for students with various physical attributes is minimal and inconclusive. Rogers et al. (2012) found that 19 of the 48 studies reviewed focused on students with either visual or hearing impairments. For students with visual impairments, accommodations included, but were not limited to, audio recordings, computerized oral presentations, large print, and read-aloud, with oral presentations and computer-administrated accommodations showing positive effects on test scores. Moreover, use of large print, braille, or read aloud accommodations were shown not to alter the construct being assessed. Cawthon's research (2009, 2010) on the prevalence of accommodation use with students who are deaf or hard of hearing was also summarized by Rogers et al. (2012). They reported that small group administration, extended time and interpreting test directions as being used most frequently with this population. Effects of these accommodations for students with hearing impairments were not reported, nor were effects of accommodations for students with physical disabilities. However, in an earlier review, Cormier et al. (2010) reported that of the 30 studies published during 2007–2008 including a student sample, five studies focused on students with vision or hearing impairments or other physical disabilities. The same authors also noted that accommodated tests for students with physical disabilities were not highly comparable with test scores of non-accommodated students without physical disabilities.

We designed the SAAT to gather information about these personal characteristics, or others, to assess a student's vulnerability to CIV. Additional research on the effectiveness of accommodations for mediating the impact of students' personal characteristics on CIV is needed.

Steps 3-4: Test and Item Specifications for the SAAT

When designing a test to meet a specific purpose, careful attention must be paid to the test and item specifications, including the test and item format, data collection procedures, administration, scoring, and reporting (Downing, 2006). Severson et al. (2007) recommends designing screening systems that are multi-method (i.e., integrate different sources of data), multi-agent (i.e., incorporate perspectives from various people), and multi-setting (i.e., collect information across various locations and times). We incorporated these guidelines in designing the test and item specifications for the SAAT to determine if students' scores might be compromised due to CIV.

Multi-method system of screening. We designed the test and item specifications for the SAAT to incorporate data from multiple sources. The SAAT is comprised of a series of surveys divided into three sections of academic engagement. Multiple sources of data include performance data from classroom tasks and subjective data from teachers, parents, and students on Likert-type items. The surveys are computer administered for ease in aggregating data, and reporting and interpreting results. By including various approaches, including rating scales and direct skills measurement, accuracy of screening may be improved (Elliott et al., 2007).

Performance data are collected from classroom tasks in the areas of reading, writing, and mathematics (application and computation). This information is used to point to areas of the curriculum that might interfere with the student's ability to demonstrate his or her mathematics knowledge and skills. For example, if a student has difficulty responding to literal reading comprehension questions, he or she may have difficulty retaining story information

in working memory. When responding to a mathematics test, a student with these difficulties may not be able to retain and sort relevant information from irrelevant information when solving multi-step word problems. When collecting student performance data, multiple sources of instructionally relevant data should be included (i.e., standardized assessments, classroom assessments, in-class work, homework). Observational data can also be incorporated to verify assessment results. Regardless of the performance measures that are used, data are entered into the SAAT and then aggregated. To arrive at the summary judgment displayed on the Summary Report described below, the teacher or IEP team reviews the performance data and comes to consensus on the reporting category. If appropriate, data used to generate the present level of performance from the Individualized Education Program (IEP) can be entered. Performance data are integrated with subjective data sources on the subsequent reports.

Subjective data sources include perceptions from teachers, parents, as well as students, and support a multi-method approach. These data may provide valuable insights needed to interpret student performance data to determine possible risk factors that might lead to CIV. Moreover, in a summary of the accuracy of teachers' judgment in making screening decisions, Elliot et al. (2007) noted that when data are collected in a systematic and structured manner, teachers' accurately identify students who are experiencing difficulties. Also, collecting perception data is often more cost effective and time efficient than gathering performance data.

The SAAT collects subjective data on three dimensions of a student's academic engagement: perception of proficiency in academic tasks, possible barriers to academic success, and possible solutions to overcome barriers (see Figure 2 for a sample of the Teacher Survey). These surveys are modeled after previously developed surveys that have been reviewed and evaluated by educators (Alonzo, Ketterlin-Geller, & Tindal, 2004).

In Section 1, data are collected from teachers, parents, and the student on their perceptions of the student's proficiency in academic tasks. These data are triangulated with the student's classroom performance data to determine correspondence between these sources. Because the student's performance on classroom tasks may be confounded with unidentified sources of CIV, we collect perception data to verify the performance data. Data are collected on a 4-point scale from *not proficient* to *highly proficient*.

In Section 2, teacher, parent, and student perceptions about possible barriers to academic success are collected. The surveys gather data that are aligned with aspects of academic engagement that are associated with the risk factors that may lead to CIV. For example, perceptions about the student's sensory engagement are solicited by asking if it is difficult for the student to "take a test in a room with background noise" or "take a test with multiple questions per page." The perceived difficulties in engaging with these school-related tasks are captured on a 4-point scale from *not difficult* to *very difficult.*

In the final section, Section 3, the perception surveys collect data on possible solutions to overcome the barriers previously identified. Because teachers, parents, and the individual student often have devised solutions to increase engagement with academic tasks, we integrated a mechanism for capturing this information. The amount of help a specific solution may provide is solicited from teachers, parents, and the student on a 4-point scale (1 = *no help* to 4 = *a lot of help*). These data may point to possible accommodations for the student.

Multi-agent and multi-setting assessment. Multiple people interact with and provide educational support to students with disabilities across various settings including school and home. Each interaction provides additional information about a student's needs. To incorporate these perspectives in an efficient manner, we designed the test and item specifications for the SAAT to purposefully gather input from multiple informants including the special education teacher, general education teacher, parents, and student.

Limited research is available on teachers', parents', and students' perceptions of test accommodations. However, some findings point to possible misconceptions about accommodations, which may introduce sources of bias when assigning accommodations. In a survey collected from 98 teachers in Ontario, Canada, Brackenreed (2004) found that the type of impairment impacted the teachers' perception of acceptability of the accommodation. Although most teachers agreed that response-mode accommodations for students with sensory impairments did not change the tested construct, they perceived accommodations that changed the test format and administration (such as extended time, reducing the number of items per page, and rewording questions) as changing the tested construct. In a similar study with 35 teachers and 43 parents, Lang et al. (2005) found that most participants viewed accommodated tests taken by students with disabilities as only somewhat comparable to results from non-accommodated tests taken by students without disabilities. In the event that teachers' and parents' prior perceptions of the appropriateness of accommodations influence their willingness to recommend accommodations, bias may be introduced in the accommodation assignment process. Instead of directly probing about accommodations, the SAAT collects data from teachers and parents about students' personal

Please read each statement and mark the description that best aligns with this student.

Section 1 – Student's Demonstrated Proficiency					Key
Rate the student's proficiency in:	Not proficient ○	Emerging proficiency ◐	Proficient ◑	Highly proficient ●	
Decoding grade-level text					L
Comprehending grade-level text					L
Writing to express oneself					L
Solving mathematics word problems					M
Solving math computation problems					M
Section 2 – Possible Barriers to Academic Success					**Key**
How difficult is it for the student to:	Not difficult ○	Sort of difficult ◐	Difficult ◑	Very difficult ●	
Remember information learned recently					M
Remember information learned over time					M
Read and understand written directions					L
Read and understand mathematics word problems					L
Understand the meaning of words in mathematics word problems					L
Express one's thinking in numbers, words, or phrases					L
Handwrite or type answers to test questions					P
See the numbers, letters, or pictures in test questions					P
Use a computer					P
Use a calculator to solve math problems					P
Complete a test within the allowed amount of time					P
Work independently for 45-60 minutes					S
Take a test in a room with background noise					S
Take a test in a room with other people or distractions					S
Take a test with multiple questions per page					S
Section 3 – Possible Solutions to Overcome Barriers					**Key**
How much would it help the student to:	No help ○	Some Help ◐	Help ◑	A Lot of Help ●	
Use a graphic organizer to solve mathematics word problems					M
Use a calculator to solve mathematics computation problems					M
Use models or objects (manipulatives) to solve math problems					M
Have directions read aloud or signed					L
Have test items read aloud or signed					L
Use a dictionary for non-mathematical vocabulary					L
Read directions or take tests that use simplified language					L
Respond to questions in alternate formats such as typing, pointing, with the use of a scribe, or other assistive devices					P
Magnify text or use Braille to read directions or take tests					P
Use a computer for taking tests					P
Take a test in multiple short testing sessions					P
Work in an environment with minimal distractions (e.g. study carrel)					P
Work alone in a separate location					P
Take a test that has been reformatted to include fewer numbers of questions per page					P
Use adaptive equipment/furniture to provide physical support Please specify: _____					P

Figure 2. Sample Teacher Survey of the Screening to Assign Accommodations Tool. This sample is not inclusive of all questions on the survey. The key on the right column is associated with common learner characteristics, and is for illustration purposes only. Key: M = Poor working or long-term memory; S = Sensory sensitivities; L = Linguistic or language difficulties; P = Physical attributes

characteristics that might serve as risk factors for CIV. We hypothesize that bias will be minimized.

Student perception of test accommodations may also impact their willingness to request accommodations during IEP meetings. In a study with students in grades 4 and 8, Lang et al. (2005) found that 62% of students with disabilities and 50% of students without disabilities classified the accommodated test as easier than the non-accommodated test. Similarly, in a later study, Lang, Elliott, Bolt, and Kratochwill (2008) found that just under half of the students with and without disabilities preferred the accommodated condition. However, in each of these studies, most of the students with and without disabilities felt the accommodations were fair.

By designing the test and item specifications of the SAAT to collect information using a multi-method, multi-agent, and multi-setting approach, data obtained are cross referenced with other sources to improve the accuracy of the screening results. Future research is needed to examine the impact of teachers', parents', and students' perceptions about accommodations on their evaluation of possible risk factors for CIV.

Content Review of the SAAT

Once the test and item specifications were established, content-related evidence for validity was needed to verify alignment with the prototype versions of the instrument and report. Experts in special education and mathematics teachers were solicited to gather content-related evidence for validity. Reviewers were asked to provide input on the appropriateness of the content as well as feasibility and utility of the tool for making accommodations decisions. This included (a) comments on the usefulness of the information solicited for making accommodations decision, (b) perceptions about the utility of this information to support decision making, and (c) recommendations for other information that would be useful. Reviewers were given specific directions for completing the review. Eight experts and teachers agreed to participate in the external review.

Participants. Three special education experts and three mathematics teachers completed the review in the allotted time. All special education experts had an earned Ph.D. in special education or related fields. Their professional careers ranged from 5 to 40 years in areas including research, policy, and teaching focused on the inclusion and accommodation of students with disabilities in educational assessments. Mathematics teachers (n = 3) had 9 to 22 years of experience teaching mathematics in elementary, middle, and high school; all had experience supporting students with disabilities. Two reviewers submitted incomplete or late reviews; their feedback was not included.

Usefulness. When asked about the usefulness of the performance data collected in Section 1, there was general consensus among reviewers that the solicited information is useful for making accommodations decisions. All reviewers noted the value of the information for identifying instructional supports as well as broad areas in which the student needs extra academic support. One reviewer suggested linking student performance descriptions to grade level instruction; another reviewer commented that some teachers may be uncomfortable rating students' reading performance due to limited experience evaluating reading.

All reviewers reported the usefulness of knowing students' engagement levels as collected in Section 2. Reviewers noted that the information is useful for assigning accommodations and supporting student learning. Some reviewers suggested additional dimensions to evaluate such as "understand and follow oral directions" and "organize material." Several reviewers suggested defining and providing examples of the dimensions of engagement.

Section 3 solicits information about possible solutions to issues with engagement. Reviewers identified this information as useful for making accommodations decisions and supporting student learning. Some additional solutions recommended by the reviewers include "use a multiplication chart" and "use an overlay or placeholder". One expert recommended collecting evidence to verify the perceptions gathered about the usefulness of the solutions and the student's experience using each solution.

Feasibility and utility. All respondents noted that the SAAT is feasible for teachers to implement and interpret. Most respondents noted that the length was reasonable, and predicted that it would take about 5 to15 minutes to complete. One reviewer reported a completion time of 30 minutes. Directions were easy to understand, and the graduated rating scale was helpful for interpreting the responses. All reviewers complimented the visual layout of the instrument. One reviewer requested additional clarification about different ratings on the scale. An additional reviewer suggested repeating the response options at the top of each page.

Reviewers commented that the SAAT report was easy and time efficient to read and understand. Most reviewers (n = 5) could articulate the findings in the report from Sections 1, 2, and 3, and stated that the information was presented clearly. One reviewer noted some confusion when interpreting the report, but was able to understand it after rereading the material. Another reviewer recommended clarifying the rating categories and summary judgment

on the report. Finally, two reviewers recommended developing teacher training materials and/or a manual.

The content-related validity evidence collected supports the use of the SAAT for making accommodation assignment decisions. Evidence supports the test and item specification decisions, and suggests that the SAAT provides useful information. Additional evidence was collected about the accessibility and ease of use, providing supporting evidence for Step 5.

Step 5: Design and Assembly of the Test

Formatting and organizing the test items into the final protocol is an important test development step. Although highly coupled with Steps 3 and 4, designing the operational test protocol should account for the users' experiences by maximizing readability and ease of use, minimizing cues, and reducing possible sources of error. For the SAAT, the external reviewers evaluated the protocol on important dimensions of test design and assembly. Input gathered from the external reviewers was taken into account when finalizing the SAAT.

Steps 8, 10: Scoring Responses and Reporting Results to Assess an Individual's Risk for CIV

To assign meaning to the responses collected by a test, a scoring system needs to be applied (Downing, 2006). Use of test results is often predicated on providing reports that are timely, easy to interpret, and accurately display information related to the construct. Glover and Albers (2007) note that administration, scoring, and interpreting reports from screeners should be feasible for practitioners so as to increase use of the screening system.

People are naturally adept at understanding spatial information communicated through graphs, charts, and pictures (Wainer, 1992). In many instances, communication of data is enhanced through spatial representation. Today's technology makes it easy to collect and store data—more data than we can effectively use—therefore, careful organization and representation is critical. "Bronowski [1978] points out, we see the world the way *we* look rather than the way *it* looks, which constrains what we perceive" (Wainer & Velleman, 2001, p. 316). If this is the case, educational decision making might be influenced through the lens used by the individual participant to view data. The viewing and discussing of visual displays by a team is therefore critical because each team member will interpret the data according to his or her own lens.

To assess an individual's level of risk for CIV, performance data as well as data collected from teachers, parents, and the student on the SAAT are summarized in two reports: SAAT Summary Report and SAAT Risk Factors Report. The SAAT Summary Report presents information in tabular form to indicate the level of agreement across sources. The SAAT Risk Factors report integrates and synthesizes information across the various sections of the Summary Report into a visual display of the student's level of risk and possible solutions.

The tables in the SAAT Summary Report highlight trends rather than focus on a single data point, illustrate relationships between different data sources and perspectives, and bring attention to outliers. In the table, the "Summary Judgment" displays the modal score across respondents. In instances where there is discrepancy of more than one point between any two respondents, a check mark is placed in the column labeled "Topic for Discussion." These disagreements can be discussed in detail at the IEP team meeting. The left diagram in Figure 3 displays a portion of the SAAT Summary Report for one student.

The SAAT Risk Factors Report provides an integrated view of the information presented in the SAAT Summary Report. By displaying these data in tabular form as well as graphically, users can easily evaluate the student's level of proficiency, probable level of risk, and potential solutions for mediating the risk factors for CIV. Figure 3 identifies how the SAAT Risk Factor Report is generated from the SAAT Summary Report. The possible solutions identified in the report align with accommodations that may reduce CIV. Individual diagnostic information can be gathered to determine if the student receives a differential boost from these solutions.

Gathering Additional Data about Students' Needs

After screening for risk status, relevant data should be collected to guide future decision- making (Glover & Albers, 2007). For students identified as being at risk, diagnostic assessments are administered to evaluate the nature and extent of an individual's problem in order to determine the most appropriate intervention. If a student has minimal to no risk, there is little cause for using instructional time to gather diagnostic information.

Accommodations are intended to minimize the impact of CIV on students' scores. If an accommodation is successful at mediating the effect of CIV, a differential boost is observed between students' scores on an accommodated version of a test as compared to a non-accommodated version (Fuchs & Fuchs, 2001). Research conducted by Fuchs and colleagues provides evidence that an individual diagnostic approach to verifying accommodations assignment improves the accuracy of using accommodations to mediate the impact of CIV. In a study conducted by Fuchs, Fuchs, Eaton, Hamlett, and

Student Name: Zachary Robyns Grade: 3 Date Generated: May 6, 2013

Screening to Assign Accommodations Tool Summary Report

The purpose of this report is to summarize student performance data and information received from the teacher, parent, and student surveys.

Student Performance Data

Academic Performance Across Multiple Sources				
Student Performance Data	Source 1	Source 2	Source 3	Summary
Decoding grade-level text	ORF: 72 wpm	Observation 1: slow, but accurate	Observation 2: made several mistakes	Emerging Proficiency
Comprehending grade-level text	M/C Test: 12/15 correct	Maze: 5/10 correct	Class discussion: some literal, but limited inferential	Emerging Proficiency

Perception Survey Results

The symbol represented in each cell corresponds to the perception noted by the respondent.

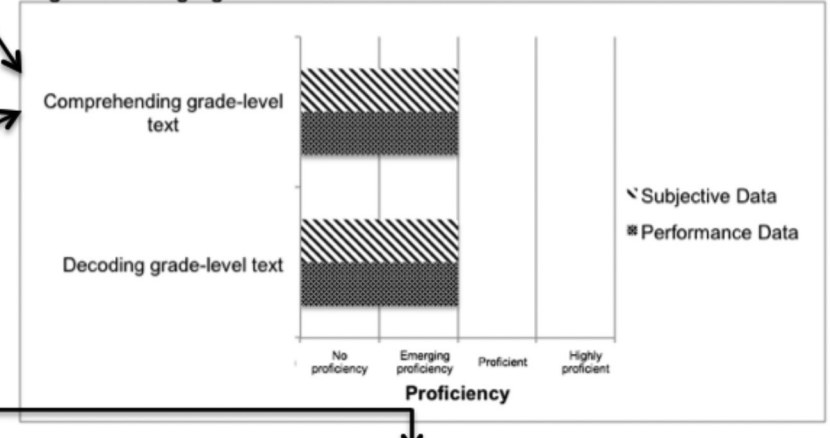
Section 1 – Student's Demonstrated Proficiency						
Rating Scale:	Not proficient	Emerging proficiency	Proficient	Highly proficient		
Rate the student's proficiency in:	Teacher 1	Teacher 2	Parent	Student	Summary Judgment	Topic for Discussion
Decoding grade-level text	○	●	○	○	○	
Comprehending grade-level text	○	○	○	●	○	
Section 2 – Possible Barriers to Academic Success						
Rating Scale	Not difficult	Sort of Difficult	Difficult	Very Difficult		
How difficult is it for the student to:	Teacher 1	Teacher 2	Parent	Student	Summary Judgment	Topic for Discussion
Read and understand written directions	●	○	●	●	●	
Read and understand mathematics word problems	●	●	●	●	●	
Understand the meaning of words in mathematics word problems	●	●	●	●	● ●	✔
Section 3 – Possible Solutions to Overcome Barriers						
Rating Scale	No Help	Some Help	Help	A Lot of Help		
How much would it help the student to:	Teacher 1	Teacher 2	Parent	Student	Summary Judgment	Topic for Discussion
Have directions read (or sign) aloud	●	●	●	●	● ●	
Have test items read (or sign) aloud	●	●	●	●	●	
Use a dictionary for non-mathematical vocabulary	●	●	○	●	●	✔
Read directions or take tests that use simplified language	●	○	○	○	○	✔

Student Name: Zachary Robyns Grade: 3 Date Generated: May 6, 2013

Screening to Assign Accommodations Tool Risk Factors Report

The purpose of this report is to integrate and synthesize student performance data and information received from the teacher, parent, and student surveys.

Linguistic or Language Difficulties



Risk Factors	Level of Risk	Potential Solutions
Read and understand written directions	Moderate risk	✓ Have directions read (or sign) aloud
Read and understand mathematics word problems	High risk	✓ Have test items read (or sign) aloud
Understand the meaning of words in mathematics word problems	Moderate-high risk	? Use a dictionary for non-mathematical vocabulary

Figure 3. Sample SAAT Summary Report and SAAT Risk Factor Report for the Screening to Assign Accommodations Tool indicating how the SAAT Risk Factor Report is generated from the Summary Report. Only information the student's linguistic or language difficulties is presented in this sample. The full report aligns with all dimensions of the SAAT.

Karns (2000), students took a series of curriculum-based measures in mathematics computation and problem solving under different accommodation conditions to determine the conditions under which each student was best able to demonstrate his or her abilities. Results from these studies provided an empirical basis for selecting the most appropriate accommodations for these individuals.

Using the results from the SAAT, teachers can design and implement an individual diagnostic approach to verify the appropriateness of the recommended accommodations. Gathering empirical data provides additional evidence as to which accommodations reduce the impact of CIV on a student's mathematics test performance.

Discussion

"[S]creening is a critical prerequisite to providing early school-based prevention and intervention services for students at risk or with… difficulties" (Glover & Albers, 2007, p. 117). Screening in education is most commonly used to identify students with academic or behavioral difficulties as part of a system-level decision-making system. In this article, we contend that screening can be appropriately applied to identify students for whom CIV might obscure accurate measurement of their mathematics ability.

Addressing the problem of inappropriate assignment of accommodations, we designed the Screening to Assign Accommodations Tool to provide IEP teams with a comprehensive information system to support appropriate assignment. Using steps from Downing's (2006) test development process and characteristics of screening systems (c.f., Glover & Albers, 2007; Elliot et al., 2007), we created a screening tool that can be used to identify students whose disabilities may interfere with accurate measurement of their mathematics knowledge, skills, and abilities leading to CIV. The SAAT takes into account students' personal characteristics when recommending accommodations in a multi-method, multi-agent, and multi-setting approach to improve the accuracy of the screening results.

The SAAT is designed to provide IEP teams with cost effective and time efficient information to make accommodation decisions. In a review of the usability and feasibility of the SAAT, reviewers noted the value of the information collected for making accommodation decisions and the importance of organizing student performance data and linking student's individual characteristics with possible accommodations in an integrated system. Also, reviewers reported that the tool was easy to understand and could be completed in about 15 minutes.

Limitations and Future Research

Although we describe in detail the test development process, limitations should be noted. First, we solicited feedback to document initial validity evidence; however, we sampled a small population of experts, and not all reviewers who agreed to participate submitted feedback. Therefore, prior to using this tool in a pilot study, we should consider seeking additional input on the tool for further refinement. Second, to date the SAAT has not been evaluated for accuracy of accommodation recommendations. Future research is needed to corroborate the recommended accommodations. Third, as with other screeners, misidentification of individual's risk status can be a significant problem (Albers, Glover, & Kratochwill, 2007); research studies are needed to examine and address misidentification. Finally, the consequential validity of using the SAAT to make accommodations decisions should be examined. Simply identifying students who may be at risk for inaccurate measurement will not ensure that students receive appropriate accommodations; teachers may need targeted professional development focused on how to implement accommodations with fidelity.

In conclusion, to measure a student's progress in the general curriculum as required by the Individuals with Disabilities Education Improvement Act of 2004 (IDEIA, 2004), we must ensure that assessment results accurately reflect the knowledge and skills of all students, including students with disabilities whose disabilities may manifest in ways that interfere with accurate measurement of their knowledge, skills, and abilities in mathematics—in the absence of appropriate accommodations. Implementing appropriate accommodations, as determined by the IEP team, is critical to reducing irrelevant errors and making accurate inferences about a student's knowledge and skills. With the importance of valid measurement in mind, the SAAT was developed to assist IEP teams in making accurate, individualized accommodation decisions.

References

Ahn, R. R., Miller, L., Milberger, S., & McIntosh, D. N. (2004). Prevalence of parents' perceptions of sensory processing disorders among kindergarten children. *American Journal of Occupational Therapy, 58*(3), 287–293.

Albers, C. A., Glover, T. A., & Kratochwill, T. R. (2007). Where are we, and where do we go now? Universal screening for enhanced educational and mental health outcomes. *Journal of School Psychology, 45*, 257–263.

Alonzo, J., Ketterlin-Geller, L. R., & Tindal, G. (2004). *Instrument Development: Examining the Appropriateness of Student and Teacher Surveys for*

Determining the Need for Testing Accommodations (Tech. Rep. No. 31). Eugene, Oregon: University of Oregon, College of Education, Behavioral Research and Teaching.

Berch, D. B., & Mazzocco, M. M. (2007). *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities.* Baltimore, MA: Brookes.

Brackenreed, D. (2004). Teachers' perceptions of the effects of testing accommodations. *Exceptionality Education Canada, 14*(1), 5–22.

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

Crawford, L., & Ketterlin-Geller, L. (2013). Middle school teachers' assignment of test accommodations. *The Teacher Educator*, *48(1)*, 29–45.

Downing, S. M. (2006). Twelve steps for effective test development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 3–25). Mahwah, NJ: Lawrence Erlbaum Associates.

Edgemon, E. A., Jablonski, B. R., & Lloyd, J. W. (2006). Large-scale assessments: A teacher's guide to making decisions about accommodations. *Teaching Exceptional Children, 38*(3), 6–11.

Elbaum, B., Arguelles, M. E., Campbell, Y., & Saleh, M. B. (2004). Effects of a student-read-aloud accommodation on the performance of students with and without learning disabilities on a test of reading comprehension. *Exceptionality, 12*, 71–87.

Elliott, S. N., Huai, N., & Roach, A. T. (2007). Universal and early screening for educational difficulties: Current and future approaches. *Journal of School Psychology, 45*, 137–161.

Fuchs, L. S., & Fuchs, D. (2001). Helping teachers formulate sound test accommodation decisions for students with learning disabilities. *Learning Disabilities Research and Practice, 16*(3), 174–181.

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*(1), 67–81.

Fuchs, L. Fuchs, D., Eaton, S. B., Hamlett, C. L., & Karns, M. (2000). Supplementing teacher judgments of mathematics test accommodations with objective data sources. *School Psychology Review*, *29*(1), 65–85.

Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2007). Strategy Use, Long-Term Memory, and Working Memory Capacity. In D. B. Berch & M. Mazzocco, (Eds.), *Why is math so hard for some children?* (pp. 431–440). Baltimore, MA: Brookes.

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research, 79*(3), 1202–1242.

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117–135.

Haladyna, T., & Downing, S. M. (2004). Construct irrelevant variance in high stakes testing. *Educational Measurement: Issues and Practice*, *23*(1), 6–16.

Ives, B. (2007). Graphic organizers applied to secondary algebra instruction for students with learning disabilities. *Learning Disabilities Research & Practice, 22*(2), 110–118.

Jitendra, A. K., Star, J. R., Starosta, K., Leh, J. M., Sood, S., Caskie, G., Hughes, C. L., & Mack, T. R. (2009). Improving seventh grade students' learning of ratio and proportion: The role of schema-based instruction. *Contemporary Educational Psychology, 34*, 250–264.

Ketterlin-Geller, L. R., Alonzo, J., Braun-Monegan, J., & Tindal, G. (2007). Recommendations for accommodations: Implications of (in)consistency. *Remedial and Special Education, 28*(4), 194–206.

Ketterlin-Geller, L. R., Yovanoff, P., & Tindal, G. (2007). Developing a new paradigm for accommodations research. *Exceptional Children, 73*(3), 331–347.

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*(1), 107–124.

Lang, S. C., Kumke, P. J., Ray, C. E., Cowell, E. L., Elliott, S. N., Kratochwill, T. R., & Bolt, D. M. (2005). Consequences of using testing accommodations: Student, teacher, and parent perceptions of and reactions to testing accommodations. *Assessment for Effective Intervention, 31*(1), 49-62.

Maccini, P., & Gagnon, J. (2006). Mathematics instructional practices and assessment accommodations special and general educators. *Exceptional Children*, *72*(2), 217-234.

Moschkovich, J. N. (2010). *Language and mathematics education: Multiple perspectives and directions for research.* Charlotte, NC: IAP Information Age Publishing.

Partnership for Assessment of Readiness for College and Careers. (2013). *Assessment administration guidance.* Retrieved from http://www.parcconline.org

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). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Santoli, S. P., Sachs, J., Romey, E. A., & McClurg, S. (2008). A successful formula for middle school inclusion: Collaboration, time, and administrative support. *Research in Middle Level Education Online, 32*(2), 1–13.

Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *Journal of School Psychology, 45*, 193–223.

Smarter Balanced Assessments. (2012). *Supporting Document: Scoring Reporting and Estimated Testing Times.* Retrieved from http://www.smarterbalanced.org/smarter-balanced-assessments/

Talay-Ongan, A., & Wood, K. (2000). Unusual sensory sensitivities in Autism: A possible crossroads. *International Journal of Disability, Development & Education, 47*, 201–212.

Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher, 21*(1), 14–23.

Wainer, H., & Velleman, P. F. (2001). Statistical graphics: Mapping the pathways of science. *Annual Review of Psychology, 52*, 305–335.

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

Leanne R. Ketterlin-Geller is an associate professor of education policy and leadership at Southern Methodist University.

Lindy Crawford is an associate professor of special education at Texas Christian University.

Jacqueline N. Huscroft-D'Angelo is a research associate at Texas Christian University.

Please send correspondence to Leanne R. Ketterlin-Geller, lkgeller@smu.edu