The Gordon Commission
on the Future of Assessment in Education

Accommodation for Challenge, Diversity, and Variance in Human Characteristics

Martha L. Thurlow
National Center on Educational Outcomes
University of Minnesota

Assessment accommodations emerged during the 1990s as the avenue to access for students with disabilities. Defined at that time as changes in materials or procedures that provided access to large-scale assessments, accommodations became both the hope of students, parents, and educators of students with disabilities, and the bane of test developers and psychometricians.

Accommodations were seen as the way to ensure that students with disabilities could participate in national, state, and district assessments at a time when lack of participation was the norm rather than the exception (cf. Kantrowitz & Springen, 1997; Shriner, Spande, & Thurlow, 1993; Ysseldyke, Algozzine, & Thurlow, 1992; Ysseldyke, Thurlow, & Linn, 1999; Zlatos, 1994). Today, accommodations are viewed as a way to ensure that assessments really avoid measuring construct irrelevant variance and instead measure what students with disabilities know and are able to do. Accommodations are defined in a variety of ways that conform to this critical element – the validity of results when accommodations are used (cf. Bolt & Roach, 2009; Laitusis & Cook, 2007).

Much has changed since the early 1990s, with nearly 99% of students with disabilities participating in state assessments. Just as participation rates and perceptions of accommodations have changed, so too has the context of instructional and assessment systems. These changes are part of an evolution that leads to continued questions about the extent to which accommodations are needed, and who needs them.

The purpose of this chapter is to explore some of the continuing evolution in instructional and assessment accommodations. I do this by providing background on the theoretical perspectives underlying assessment accommodations, including the history of accommodations, accommodation policies, and validity considerations. I then address changing perspectives on

accessible assessments and the implications of these changes for instructional and assessment accommodations. I also address the question of "who needs accommodations?" in an attempt to explore the relationship been the continuum of student characteristics and the need for accommodations. Finally, I summarize lessons learned from disability research and practice that have application to students other than those with identified disabilities or those with limited English proficiency. I do not attempt to address accommodations in the array of alternate assessments that federal policy has allowed for students with disabilities.

Theoretical Perspectives

A landmark study by the National Research Council (McDonnell et al., 1997) confirmed the importance of the participation of students with disabilities in large-scale assessments, particularly statewide assessments, if students with disabilities were to benefit from standards-based educational reforms. This study also pointed to the importance of accommodations for students to have access to the general curriculum, and to be able to show their knowledge and skills on assessments.

Perspectives on assessment accommodations have changed quite dramatically over time. Initially, accommodations were viewed as a way to level the playing field as well as to improve the validity of assessment results (Burns, 1998; McDonnell, McLaughlin, & Morison, 1997; Thurlow, Elliott, & Ysseldyke, 1998) and to provide a way for students to take assessments when they otherwise would not be able to take them. These goals do not necessarily mean that student performance will improve (Thurlow, Elliott, & Ysseldyke, 2003). Accommodations also were viewed as an important part of instruction for students with disabilities. But, instructional accommodations covered changes such as modifying the number of items in a classroom assignment or including only easier parts of an assignment — changes well beyond those

considered for assessment (Elliott & Thurlow, 2006; Nolet & McLaughlin, 2005). Despite the recommendation that only accommodations used in instruction should be used for assessment, concern grew about making changes that actually altered what a test was intended to measure.

As policy and practice evolved, so did the language surrounding accommodations. With time, those changes that produced valid results were called accommodations, whereas those that altered what the test was intended to measure were called modifications.

Almost exclusively, the early belief was that accommodations were reasonable to provide only to individuals with disabilities, and even then, there were many concerns about whether accommodations provided an advantage to those students with disabilities using them (Koretz & Barron, 1998; Koretz & Hamilton, 2000). Shifts in these early perspectives occurred over time with the enactment of policy imperatives for students with disabilities to be included in assessments, to the point that the discussion focused on the disadvantage imposed on students who did not receive needed accommodations:

> Accommodations are designed to decrease noise and maximize the strength of the inference based on a student's score. Without accommodations, many students are at a disadvantage in demonstrating what they actually know and can do. (Madaus, Russell, & Higgins, 2009, p. 182)

History

The first academic treatment of assessment accommodations occurred in the 1980s by testing companies focused on admission tests for higher education. The American College Testing (ACT®) program and the Educational Testing Service (ETS®) each engaged in documentation of eligibility and accommodation concerns. ACT required that individuals with disabilities have a professional diagnosis and appropriate documentation of a disability before

being allowed to take the ACT assessment under "nonstandard" conditions. It was the first company to gather and analyze data from those individuals who used one or more of the allowed accommodations, which included extended time, magnifying glass, large type, Braille edition, Brailler, audio cassette editions, use of a reader, assistance filling out the answer sheet, slate and stylus, use of a tape recorder, and signing of instructions (see Thurlow, Ysseldyke, & Silverstein, 1995).

Laing and Farmer (1984) used data from five years of ACT administrations (1978–79 through 1982–84) to examine the predictive validity of ACT results for students' GPAs during their first year of college. They grouped individuals into five groups: (a) individuals with no disabilities, (b) individuals with disabilities who had a standard administration (i.e., no accommodations), (c) individuals with visual impairments who had a nonstandard administration (i.e., accommodations), (d) individuals with hearing impairments who had a nonstandard administration, and (e) individuals with "motor" disabilities (which included physical and learning disabilities) who had a nonstandard administration. No analyses could be conducted for the individuals with hearing disabilities because fewer than nine were included in the data. Laing and Farmer found that under standard conditions, the ACT was equally predictive for both those with and those without disabilities. Under nonstandard conditions, predictions were close for those with visual impairments but less for students with physical and learning disabilities.

ETS had similar findings when it examined standard and nonstandard versions of the Scholastic Aptitude Test (SAT®) and the Graduate Record Examinations (GRE®) (Willingham, Ragosta, Bennett, Braun, Rock, & Powers, 1988). The ETS researchers investigated several specific aspects of comparability (score comparability — reliability, factor structure, differential item functioning, prediction of performance, admissions decisions; task comparability — test

content, testing accommodations, test timing), and as a result had many specific findings. For example, looking at mean performance and reliability, Bennett, Rock, and Kaplan (1985) found that mean SAT scores and reliability for individuals with visual impairments were comparable to those of a national sample. Bennett, Rock, and Jirle (1986) found that mean GRE analytical scores for students with visual impairments who received large type and extended time were above those of a national sample (with test completion slightly higher for students with visual impairments than for a national sample). For students with physical disabilities, Bennett et al. (1985) found that with extended time the mean SAT scores and reliability were comparable to those of a national sample.

This early research was ground-breaking in many ways, yet it also suffered from many complexities that confounded results. Many hundreds of studies have been conducted since these early studies, and they have continued to portray the complexities of conducting research on accommodations and reaching strong conclusions based on the research (see, for example, Thurlow, Lazarus, & Christensen, 2013). I will highlight these complexities later when I describe both the approaches that researchers take to examine accommodations and their effects, and when I highlight recent research on the effects of accommodations. Suffice it to say for now that the early research began to wake policymakers to the notion of accommodations in a broader sense than simply as a mechanism for taking college entrance exams.

Early research confirmed that the availability of test accommodations was one of the keys to ensuring greater large-scale test participation, not just by students with disabilities, but also by students learning English (Mazzeo, Carlson, Voekl, & Lutkus, 2000). Because some students who were learning English also were students with disabilities, a focus on these students and

accommodations for them emerged as well (Christensen, Albus, & Thurlow, in press; Christensen, Liu, & Thurlow, 2010; Liu, Albus, & Thurlow, 2006; Thurlow & Liu, 2001).

English language learners (ELLs) initially seemed to be offered only those accommodations that had been designated for students with disabilities (e.g., large print books or extra time and breaks). Significant progress has been made since then in identifying the types of accommodations that are likely to address the linguistic needs of ELLs (e.g., bilingual dictionaries, glossaries, translations). There also have been significant increases in the research conducted to determine which accommodations are appropriate (given the content assessed) and effective for ELLs (Abedi, 2004; Abedi, Hofstetter, & Lord, 2004; Kopriva, Emick, Hipolito-Delgado, & Cameron, 2007; Pennock-Roman & Rivera, 2011; Rivera, Collum, Willner, & Sia, 2006), as well as for ELLs with disabilities.

When initial conversations about the importance of accommodations occurred, the discussion was about ways to ensure that students could participate in assessments so that educators would have measures of their academic achievement similar to those obtained for the peers of students with disabilities. The conversation also focused on needing to level the playing field for these students, so that they had a fair opportunity to show what they knew and could do on state assessments. Knowing what is a "level playing field" and how to be sure that changes that are provided in the name of accommodations do not actually create an "advantage" for the students who received them emerged again and again. The critical discussion was about the *validity* of assessments when accommodations were used – were the users of assessment results getting the same information for those who used accommodations as they did for those who did not? In other words, were the assessment results valid or did the accommodations alter in some fundamental way what the assessment intended to measure?

Policy Pushing Practice

Policymaking on the participation of students with disabilities in large-scale assessments, and thus the need to provide accommodations dramatically pushed forward accommodations discussion and practice in the early 1990s and 2000s. In 1992, the first examination of state accommodations policies for state assessments indicated that only 21 states had such policies (Thurlow, Ysseldyke, & Silverstein, 1993). The policies were organized by the authors into five categories: alternate presentation, alternate response, alternate setting, alternate scheduling/time, and other. Over the years, the number of states with accommodations policies has increased to all states — 21 states in 1993 (Thurlow, Shriner, & Ysseldyke, 1994), 38 states in 1995 (Thurlow, Scott, & Ysseldyke, 1995), 39 in 1997 (Thurlow, Seyfarth, Scott, & Ysseldyke, 1997), 48 in 1999 (Thurlow, House, Boys, Scott, & Ysseldyke, 2000), and all 50 states in 2001 (Thurlow, Lazarus, Thompson, & Robey, 2002). Since that time, regular updates to participation and accommodations policies showed that all 50 states (and many of the U.S. territories) have written accommodations policies that they change and update regularly (see Christensen, Braam et al., 2011; Christensen, Lazarus, Lail, Crone, & Thurlow, 2008; Clapper, Morse, Lazarus, Thompson, & Thurlow, 2005; Lazarus, Thurlow, Lail, Eisenbraun, & Kato, 2006).

The initial push for and development of policies were clearly focused on individuals with disabilities. Yet, it was soon recognized that many of the arguments being made about the need for students to be in large-scale assessments and to have the accommodations needed to allow them to show their knowledge and skills applied to students who were "limited English proficient" (LEP — now referred to as English Language Learners, ELLs, or English Learners, ELs) (Abedi, Lord, Hofstetter, & Baker, 2000; Koenig & Bachman, 2004).

A clear change in the language in policies occurred across time. This change in language in which "accommodations" were distinguished from "modifications" now is reflected in nearly all state policies and guidelines on accommodations for K–12 state assessments (Christensen, Braam, Scullin, & Thurlow, 2011). It reflects an emphasis on the validity of results when accommodations are used.

Validity

Validity is defined in the 1999 *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, National Council on Measurement in Education: AERA, APA, & NCME, 1999) as "the degree to which accumulated evidence and theory support specific interpretations of test scores entailed by proposed uses of a test" (p.184). According to the *Standards*, evidence is accumulated to provide a basis for proposed score interpretations. The evidence may be based on (a) content, (b) response processes, (c) internal structure, (d) relations to other variables, and (e) consequences of testing.

Researchers studying accommodations also have attempted to define exactly what it means for an accommodation to produce valid results. For example, Hollenbeck (2002) indicated that four conditions needed to be met for an accommodation to be appropriate and produce valid results: (a) the provision of an accommodation needed to be determined on the basis of individual need; (b) the use of the accommodation should not change the construct measured by the assessment; (c) the use of the accommodation must result in similar inferences about a student's knowledge and skills; and (d) the accommodation must produce differential effects by student or group.

Prior to the delineation by Hollenbeck (2002), Phillips (1994) outlined three conditions that needed to be checked before allowing accommodations: (a) the accommodation should not alter what the test is measuring; (b) scores between students who received the accommodation and those who did not should be comparable; and (c) accommodations should benefit only students with disabilities, not those without disabilities. This latter condition has sometimes been referred to as an interaction hypothesis – it argued that an accommodation produces no change in scores for test takers, unless they have a disability. The interaction hypothesis has been a theoretical basis for much of the early research into the effects of accommodations (Sireci, Scarpati, & Li, 2005).

As noted by researchers documenting the research paradigms for studying the effects of accommodations (Laitusis, 2007), the interaction hypothesis was replaced by the differential boost hypothesis (Fuchs & Fuchs, 1999; Sireci et al., 2005). This theoretical perspective argues that students without disabilities might benefit (i.e., show improved performance) from the use of accommodations, but that the benefit obtained by students with disabilities would be differentially (and significantly) larger than that obtained by students without disabilities. The differential boost hypothesis has been the primary driving force in experimental and quasi-experimental studies on the effects of accommodations. Note that this theoretical basis focuses on groups of students defined by a label, but could just as well apply to students defined by a characteristic (e.g., poor decoding skills). I will discuss this more later.

Not all studies of the effects of accommodations are experimental or quasi-experimental studies. In fact, because so much data exist as a result of the yearly administration of assessments of achievement, it is possible to conduct extant data analyses that focus on the effects of accommodations. These studies typically have examined the factor structure of the assessment

when accommodations are used compared to when accommodations are not used, an approach suggested by the *Standards* (AERA, APA, & NCME, 1999) for examining the internal structure of a test as part of the validation process (see, for examples of this approach, Cook, Eignor, Steinberg, Sawaki, & Cline, 2009; Kim, Schneider, & Siskind, 2009). An appropriate accommodation results in a test with the same factor structure as the test shows when accommodations are not used. When data from large numbers of test takers are available, a Differential Item Functioning (DIF) analysis has been used as well to determine whether the use of accommodations alters the characteristics of items (see, for examples of this approach, Abedi, Leon, & Kao, 2008b; Bolt & Ysseldyke, 2006; Cahalan-Laitusis, Morgon, Bridgeman, Zanna, & Stone, 2007; Finch, Barton, & Meyer, 2009).

More recent examinations also have looked at differential distractor functioning. These studies examine whether the provision of accommodations or disability category alters the selection of incorrect responses by those who do not choose the right answer on a multiple choice item (see as examples, Abedi, Leon, & Kao, 2008a; Kato, Moen, & Thurlow, 2009).

Thus, research generally has either involved analyses of large-scale data bases or has involved testing students under various conditions. Most of the testing of students, although not all, has been done to make conclusions about groups of students. Single subject research has been used occasionally to study the effects of accommodations for individual students (see as an example, McKevitt, Marquart, Mroch, Schulte, Elliott, & Kratochwill, 2000).

Despite the increase in number of studies on the effects of accommodations using the full array of possible validity approaches recommended in the *Standards*, the results of validity research remain inconclusive for the most part. Some researchers have concluded that only extended time has a clear research basis (see, for example, Sireci et al., 2005). Although

continued research is considered to be important for demonstrating the validity of accommodated assessment results, the research likely will need to target accommodations for specific tests and for students with specific characteristics and access needs.

Theoretical perspectives have moved the field over time in a direction that emphasizes the value of experimental and quasi-experimental approaches to examine the effects of accommodations. Relatively few studies have examined instructional accommodations; most of the studies that have examined instructional accommodation have focused on secondary education settings (Gregg, 2011). Also, as becomes evident as I discuss changing perspectives in the next section, the theoretical perspectives that underlie assessment accommodations for students with disabilities have expanded to also address the needs of ELLs and to better define the characteristics of students with disabilities who benefit from accommodations.

Changing Perspectives on Accessible Assessments

Accommodations provide a way to ensure that the student has access to the assessment. As a result, considerable attention has been devoted to determining which accommodations are allowed and which are not. The policies surrounding accommodations have become complex, in part to ensure that accommodations result in a score that is more valid for a student than would be obtained if the student had not used accommodations.

As the nature of assessments has changed, so has the thinking about accommodations and what it means to have an accessible assessment. This has become most evident in the shift to computer-based assessments. Finally, the notion of universal design is an important part of what makes for an accessible assessment. The evolution of "universal design" as a concept applied to assessments provides a picture of the future of accessibility and accommodations in assessments.

Allowed Accommodations

Policies that indicate which accommodations are allowed or prohibited for state assessments have become increasingly complex over time, and vary depending on the accommodation and the state (see Table 1 for the number of states allowing and prohibiting a sample of accommodations in their 2009 policies). Policies also define which accommodations will result in lower scores if used and which scores will not be added to other scores for school accountability (reflected in Table 1 under Allowed with Restrictions column). Assessment policies seem to be driving the provision of instructional accommodations, even though most state policies indicate that students must have had access to accommodations during instruction before they are provided for assessments.

The implementation of accommodations during instruction and assessments is a challenge for educators. Several studies have documented the inconsistencies in accommodations used during assessment and during instruction (DeStefano, Shriner, & Lloyd, 2001; Hollenbeck, 2002; Ketterlin-Geller, Alonzo, Braun-Monegan, & Tindal, 2007; Rhode Island Department of Education, 2003; Shriner & DeStefano, 2003). Research confirmed that accommodations decisions made by IEP teams do not necessarily get transferred to either instructional or testing environments (Fuchs & Fuchs, 1999; Hollenbeck, Tindal, & Almond, 1998; Ketterlin-Geller et al., 2007).

Increased training on making decisions about accommodations (Christensen, Carver, VanDeZande, & Lazarus, 2011; Thurlow, Albus, & Christensen, 2009) and increased monitoring of accommodations (Christensen, Thurlow, & Wang, 2009) are ways that states and districts have attempted to improve the implementation of accommodations for instruction and assessments. These represent a significant change from the early 1990s, when state provided

virtually no support to districts and schools on making decision about and implementing

accommodations (Langley & Olsen, 2003).

Among the most controversial accommodations historically have been those that require

a human to provide the accommodations. Readers, sign language interpreters, and scribes are

among these controversial accommodations. The controversy lies, in part, in the concern that the

human will divulge the answers to questions or provide other hints about answers, even if

unintended. There also have been concerns about the security of items when humans have access

to the tests before the administration of the test. Increasingly, states are providing guidelines

about the specific ways to provide accommodations that involve humans (Christensen, Braam, et

al., 2011; Clapper, Morse, Thompson, & Thurlow, 2005). Continued concern about these types

of accommodations (see: Hodgson, Lazarus, Price, Altman, & Thurlow, 2012) has led many to

believe that the answer lies in incorporating these accommodations within computer-based

delivery systems, thereby eliminating the human element.

Online and Computer-Based Assessments

Increasing numbers of states and districts have adopted online and computer-based

assessments (Thurlow, Lazarus, Albus, & Hodgson, 2010). It is believed that technology

platforms not only will transform the efficiency of testing (Madaus et al., 2009), but it also will

open up the possibility that more accommodations can be presented via the technology platform

rather than via a human (Russell, 2011). In fact, researchers and test developers argue that more

accommodations can be part of the assessment itself — embedded features of the assessment

(NCEO, 2011) — rather than external changes in materials or procedures (Meyen, Pogio, Seok,

& Smith, 2006; Russell, Hoffman, & Higgins, 2009b; Thompson, Quenemoen, & Thurlow,

2006).

It is clear that states are moving toward computer-based testing, just as districts have done in their use of computer-based formative and interim assessments. In their 2010 review, Thurlow et al. found that 22 states had their state assessments or End-of-Course assessments presented via the computer. In a few of these states, the computer-based tests were administered only in certain grades, such as middle school or high school; most states provided them in elementary through high school. Still, the ways in which these states addressed accommodations varied. For example, all of the states allowed for the test to be provided in a paper format rather than on the computer. At the time of the review, none of the states offered a computer-based version for those students needing Braille, although since that time, Oregon has implemented this option (Shah, 2012).

The use of computer-based testing will increase dramatically when states implement common assessments supported by the Race to the Top federal funding for consortia of states to develop assessments for use in all of the states in the consortium (Center for K–12 Assessment and Performance Management at ETS, 2012). These assessment development efforts indicate that technology platforms will make it easier to address the accessibility challenges that assessments have faced in the past.

Several studies have specifically examined whether changes that were previously provided as external accommodations could be provided via the computer (Dolan, Hall, Banerjee, Chun, & Strangman, 2005; Erdogan, 2008; Kamei-Hannan, 2008; Russell, Hoffman, & Higgins, 2009a). Other studies have explored the comparability of computer-based and paper assessments, though not always specifically considering the accommodations aspect of comparability (Kingston, 2009; Russell, 1999). The pending incorporation of what were previously viewed as accommodations into technology platforms raises a host of possibilities for

greater access (see, for example, Madaus et al., 2009) as well as a number of new issues to be addressed (Thurlow, Lazarus et al., 2010; Thurlow, Quenemoen, & Lazarus, 2011).

Universal Design and Accessible Assessments

Researchers and advocates who pushed for better accommodations policy and practice also recognized that there were ways in which assessments could be improved so that they did not set up irrelevant barriers to students being able to demonstrate their knowledge and skills. The first push came through the topic of "universal design," a broad concept that covered everything from the definition of the population to be included in the assessment to specific descriptions of legible text (see Table 2, which shows the elements of universal design first described by Thompson, Thurlow, & Malouf, 2004; see also, Johnstone, Altman, & Thurlow, 2006). Others have approached the notion of universal design starting from the instructional perspective and then applying it to classroom and other assessments (CAST, 2009; Cortiella, 2009); this approach tends to address multiple means of representation, action/expression, and engagement.

States and test developers have embraced the idea of universal design. For example, states now regularly include universal design requirements in their requests for proposals to develop items and tests (Altman, Lazarus, Thurlow, Quenemoen, Cuthbert, & Cormier, 2008). The development process involves, for example, the conceptualization and construction of items and the test, expert reviews, think-aloud methods as part of a pilot or field test, and the training of test developers in the elements of universal design.

As the need for, and characteristics of, universally designed assessments were further defined, a distinction was made between access skills and target skills (Ketterlin-Geller, 2008). Target skills, as defined by Ketterlin-Geller, "encompass the cognitive engagement and content

of the construct that the test is designed to measure. These are skills about which interpretations will be made" (pp. 7–8). She defined access skills to "include levels of cognitive engagement and content knowledge and/or skills that are needed to *access* the target construct but are not the intended object of measures (Bielinski, Sheinker, & Ysseldyke, 2003) (p. 8). Kettler (2021) provided several specific examples of target and access skills: for example, a target skill of reading fluency maybe affected by access skills such as working memory, attention, letter recognition, and visual acuity. Defining the target skills in more detail than has been done in the past, and identifying the access skills that may get in the way of measuring those target skills has become a critical part of both universal design and the selection of appropriate accommodations for individual students.

Researchers have addressed various ways to apply universal design approaches to the development of assessments in general (Johnstone, Thompson, Miller, & Thurlow, 2008; Ketterlin-Geller, 2005; Thurlow, Johnstone, & Ketterlin-Geller, 2008; Thurlow, Johnstone, Thompson, & Case, 2008) and specifically in relation to computer-based assessment (Dolan et al., 2009; Russell, 2011). They also have conducted studies to explore the applicability of universal design principles (Almond et al., 2010; Johnstone, Liu, Altman, & Thurlow, 2007; Johnstone et al., 2008; see also Thurlow, Laitusis et al., 2009).

The implication of greater emphasis on universal design is that there may be reduced needs for external accommodations that are provided by humans. The decrease in reliance on humans is likely to increase validity.

Who Needs Accommodations?

Three federal laws support the provision of accommodations to individuals with disabilities: Section 504 of the Rehabilitation Act of 1973, the Individuals with Disabilities

Education Act (IDEA), and the Americans with Disabilities Act (ADA). Section 504 stated that

an institution receiving Federal funds must assure that:

> admissions tests are selected and administered so as best to ensure that, when a test is
>
> administered to an applicant who has a handicap that impairs sensory, manual, or
>
> speaking skills the test results accurately reflect the applicant's aptitude or achievement
>
> level or whatever other factor the test purports to measure, rather than reflecting the
>
> applicant's impaired sensory, manual, or speaking skills (except where those skills are the
>
> factors that the test purports to measure). (Section 84.42(b)(3))

This was the first indication that it was considered discriminatory for a test to reflect an

individual's disability rather than his or her knowledge and skills. Section 504 allows for the

provision of accommodations to individuals with disabilities, including those of school age

regardless of their eligibility for special education services. Students who receive

accommodations but are not receiving special education services are considered to be "504

students' who have 504 accommodation plans."

Public Law 94-142, the Education of All Handicapped Children Act, was passed in 1975

to ensure that students with disabilities receive a free, appropriate public education. This law and

its successor in 1990, the Individuals with Disabilities Education Act (IDEA), did not address the

topic of large-scale assessments of academic achievement, but accommodations were mentioned

in relation to instruction and the IEP. The reauthorization of IDEA in 1997 introduced, for the

first time, the notion of access to state and district-wide assessments and referenced

accommodations as an aspect of participation in assessments:

> (A) In general. – All children with disabilities are included in all general State and
>
> districtwide assessment programs, including assessments described under section 1111 of

the Elementary and Secondary Education Act of 1965, with appropriate accommodations

and alternate assessments where necessary and as indicated in their respective

individualized education programs. (B) Accommodation guidelines. – The State (or, in

the case of a districtwide assessment, the local educational agency) has developed

guidelines for the provision of appropriate accommodations. (20 U.S.C. Section. 612

(17))

IDEA was reauthorized in 2004, and added a requirement that states report on the number of

students receiving accommodations for state assessments.

ADA was enacted in 1990, and included requirements for accommodations and

adaptations to be made by businesses and agencies receiving federal funds:

(A) making existing facilities used by employees readily accessible to and usable by

individuals with disabilities; and (B) … acquisition or modification of equipment or

devices, appropriate adjustment or modifications of examinations, training materials or

policies, the provision of qualified readers or interpreters, and other similar

accommodations for individuals with disabilities. (42 U.S.C. 12/11, Section 101[9])

ADA was reauthorized in 2008, and added clarifications of the definition of a "disability," which

was defined in 1990 as an individual with a physical or mental impairment that substantially

limits one or more major life activities, who has a record of such an impairment, or who is

regarded as having an impairment. The clarifications included, for example, expanding the

"illustrative" list of "major life activities" to include as examples concentrating, thinking, and

communicating. Cortiella and Kaloi (2009) suggested that ADA 2008 has implications for

Section 504 and how that law may affect "children with disabilities, including learning

disabilities, as well as other conditions such as Attention-Deficit/Hyperactivity Disorder (AD/HD), Aspergers Syndrome, diabetes, asthma, and life-threatening food allergies" (p. 1).

The three federal laws (IDEA, ESEA, and ADA) clearly viewed accommodations as being for individuals with disabilities, but ESEA also recognized the need for accommodations for English language learners. States determined who could and who could not receive accommodations in their assessments. Table 3 shows the progression over time in who was, according to written policies and guidelines, and who was not considered eligible for an accommodation during state testing. It is noteworthy that some states indicated that all students were eligible for accommodations during state testing. Yet, as universal design and accessibility concepts emerged along with the belief that assessments needed to be accessible for the widest range of students, the belief grew that some of the changes previously viewed as "accommodations" needed by students with disabilities or ELLs, actually were appropriate for all individuals.

Characteristics that Lead to a Need for Accommodations

Even though the focus on accommodations in K–12 education started with a focus on students with disabilities, policymakers soon recognized that the category of a student's disability (such as learning disability – LD, speech language impairment – SLI, or intellectual disability – ID) was not a good basis for making decisions about the need for accommodations or which accommodations might be needed by an individual student (Elliott & Thurlow, 2006; Thurlow, Elliott, & Ysseldyke, 1998, 2003). Instead, the disability-related characteristics of a student (e.g., limited reading skills, poor memory, and distractibility) or the linguistic characteristics of a student (Koenig et al., 2004) were identified as being the appropriate basis for making decisions about accommodations. Table 4 provides a sample of characteristics, needs,

and possible accommodations identified in some accommodations manuals and other tools available on the web (Thurlow, Albus, & Christensen, 2009).

Needs not Specifically Related to Disability or Limited English Proficiency

Many of the needs identified in Table 4 as being the basis for providing accommodations to students with disabilities or ELLs are ones that are not limited to students in these groups. In fact, considerable research has indicated the overlap of characteristics of students with disabilities and students without disabilities (Almond et al., 2009; Cameto & Nagle, 2010; Wu, Liu, Thurlow, Lazarus, Altman, & Christian, 2012; Ysseldyke, Algozzine, Shinn, & McGue, 1982).

The policies of states that allow accommodations for all students provide additional information about what kinds of characteristics policymakers believe may be evident in students other than those with disabilities and those learning English. In 2009, 15 states allowed accommodations during state testing for "all students," most often with some type of qualifications on the use of accommodations (Christensen, Braam et al., 2011) — the states were: Arizona, California, Colorado, Kansas, Maine, Montana, New Hampshire, Oregon, Pennsylvania, Rhode Island, Tennessee, Texas, Vermont, Washington, and Wisconsin. For example, Kansas indicated that all accommodations allowed for its state test are allowed for general education students as well as those with disabilities and ELLs. Arizona indicated that "universal accommodations" are available to all students, as long as the teacher indicates the need; nevertheless, "standard accommodations" are available only to students with disabilities and ELLs. Colorado indicated that all students who need accommodations may use them for assessments, as long as the need was determined on an individual basis and documented in a formal plan. Maine stated that "students who may be considered for accommodations include,

but are not limited to, those who are ill or incapacitated in some way, who have limited English proficiency, who have an identified disability under [IDEA], who are identified as having disabilities under Section 504, or for whom classroom accommodations are necessary on a daily basis to measure academic achievement." Pennsylvania provides a list of accommodations available to all students in its manual:

> Certain accommodations are allowed for all students. These accommodations must be ongoing and documented in the student record file. All students may use extended time, preferential seating, separate setting, hospital/home testing, small grouping, keyboarding, special paper, reading windows, recorded directions, recorded verbatim sections of mathematics, student marks in test booklet, dictation for mathematics and reading (illegible writing, broken arm), reading aloud for all directions and mathematics test, marking answer at student direction, student use of stickers or highlighters, simplified directions.

Other states give variations of the examples provided here, some with less detail and others with more detail.

Lessons Learned from Disability Research and Practice

Research on the effects of accommodations for students with disabilities grew rapidly after the enactment of IDEA 97. The early work of Willingham et al. (1988) at ETS and researchers at ACT (Laing & Farmer, 1984) provided a model of the type of research that was needed (see also Thurlow et al., 2000). And, as noted earlier, experimental studies relied on the differential boost theoretical perspective to determine whether an accommodation was one that made a difference for students with disabilities or ELLs. Although it is not possible to provide a complete analysis of all the literature here, even if it only focused on that research targeting

students with disabilities, it is possible to draw some general conclusions about the nature of the research, the challenges in conducting this research, and its implications for research and practice. In addition, in this section, I will address changes that are occurring in the standards that guide testing practices. I conclude this section with a set of lessons learned that apply to all students.

Literature on Accommodations Research and its Implications

According to summaries of the disability research on accommodations conducted by Tindal and Fuchs (1999) and numerous investigators at the National Center on Educational Outcomes (Cormier, Altman, Shyyan, & Thurlow, 2010; Johnstone, Altman, Thurlow, & Thompson, 2006; Rogers, Christian, & Thurlow, 2012; Thompson, Blount, & Thurlow, 2002; Zenisky & Sireci, 2007), the number of studies on assessment accommodations has increased dramatically over the years. The number increased from just 8 studies published in 1990 to 24 published in 2010 (see **https://apps.cehd.umn.edu/nceo/accommodations/**).

In the 2009-2010 two-year period, the primary design of most of the studies was quasi-experimental (n = 16), followed by those considered to be descriptive quantitative (n = 5), descriptive qualitative (n = 5), and correlational (n = 1). Few studies were considered to be experimental (n = 2), meaning that they included random assignment of students to conditions. Most studies involved students with learning disabilities (n = 26), attention problems (n = 11), emotional-behavioral disabilities (n = 11), and visual impairments (n = 10). Most studies examined presentation accommodations such as computer administration (n = 9), Braille (n = 5), and sign language interpretation (n = 3). Two studies each examined clarifying directions, cuing, dictionary, examiner familiarity, and format.

Despite the increased research that is being conducted, and the fact that experimental and quasi-experimental studies are being conducted, there remain many issues with the research (Thurlow et al., 2013). For example, some studies have focused on students with one category of disability label (e.g., learning disabilities), even though the students with that label can vary considerably in their characteristics and their need for specific accommodations. We also know from studying students with and without disabilities that there are many overlapping characteristics that are likely to create a need for accommodations. When it is assumed that a student needs an accommodation based solely on a category of disability, and that any student without a disability label does not need an accommodation, there is a great likelihood that research findings will be confounded.

Only a few studies have clearly defined the students with disabilities in terms of the characteristics that are indicative of a specific need that might be addressed by an accommodation. For example, Laitusis (2010) further defined students into those with reading-based learning disabilities and those without disabilities. With this clearer definition of the groups included in the study, she also used a repeated measures approach, with each student taking a test with an audio presentation and without an audio presentation, in counter-balanced order. She found that students with reading-based learning disabilities "benefited differentially" more than students without disabilities. Previous research on this accommodation had produced conflicting findings, most likely due in part to the existence of the same characteristics and needs within the two groups being compared. Although the research by Laitusis focused on the audio presentation, students receiving the audio presentation also received extended time and recording answers in a test booklet rather than a bubble sheet as accommodations.

Unlike some previous experimental research, Laitusis recognized that accommodations rarely are used in isolation. Rather, when students use an accommodation it is often used in conjunction with other accommodations (Elliott, Bielinski, Thurlow, DeVito, & Hedlund, 1999). The use of more than one accommodation was recognized by McKevitt et al. (2000) in their early work using single subject designs, as well as by Fletcher et al. (2006, 2009) in their study of a bundled accommodations package that included reading of proper nouns, question stems, and answer choices as well as testing across two days. Crawford (2007) concluded a report on the value and validity of testing accommodations by recommending that there be "an expansion of the research base related to 'bundled accommodation,' including the effect of bundled, or packaged, accommodations on the construct being measured, on each other, and on individual students" (p. 14).

In addition to research on the effectiveness of accommodations, several studies have examined issues surrounding the implementation of accommodations (DeStefano et al., 2001; Fuchs & Fuchs, 1999; Hodgson et al., 2012; Rhode Island Department of Education, 2003; Shriner & DeStefano, 2003). Examination of the implementation of accommodations is important because research has found that there is a relationship between the provision of accommodations and participation in assessments (Anderson, Jenkins, & Miller, 1996; Cox, Herner, Demczyk, & Nieberding, 2006).

Not long after the enactment of IDEA requirements for students to participate in assessments, with accommodations as needed, researchers began to examine the implementation of accommodations during instruction and during assessments (DeStefano et al., 2001; Shriner & DeStefano, 2003). These and other findings of inconsistencies in the process for making decisions about accommodations (Ketterlin-Geller et al., 2007) spurred a number of training

efforts to improve accommodations decisions (e.g., Hodgson, Lazarus, & Thurlow, 2011; Thurlow, Albus, & Christensen, 2009). Logistical concerns were raised as well in response to concerns about the needs to have test administrators or other school personnel to provide accommodations during testing (Bowen & Ferrell, 2003; Hodgson et al., 2012). Although providing accommodations on computers will solve some of the complex issues surrounding the implementation of accommodations, it will not solve all of the issues, particularly those involved in decision making processes.

Accommodations in Standards that Guide Best Practice for Assessments

Accommodations have been recognized in standards and guidelines documents that direct practice for testing. The *Standards for Educational and Psychological Testing* addressed accommodations in its 1999 edition with the following statements:

> Fairness also requires that all examinees be afforded appropriate testing conditions. Careful standardization of tests and administration conditions generally helps to assure that examinees have comparable opportunity to demonstrate the abilities or attributes to be measured. In some cases, however, aspects of the testing process that pose no particular challenge for most examinees may prevent specific groups or individuals from accurately demonstrating their standing with respect to the construct of interest (e.g., due to disability or language background). (AERA/APA/NCME, p. 75)

The *Standards* go on to say that standardized procedures should be followed "unless explicit, documented accommodations have been made" (p. 75). Standard 7.7 states:

> In testing applications where the level of linguistic or reading ability is not part of the construct of interest, the linguistic or reading demands of the test should be kept to the minimum necessary for the valid assessment of the intended construct. (p. 82)

In a chapter devoted to testing individuals with disabilities, the term "accommodation" and "modification" are used as equivalent terms, and their purpose is described as being to "minimize the impact of test-taker attributes that are not relevant to the construct that is the primary focus of the assessment" (p. 101). A number of other points are made, including the importance of making individualized decisions and the need to ensure that changes that are made do not affect the psychometric qualities of the test. The lack of research at the time this edition was published was noted as well.

Much has changed since the 1999 edition of the *Standards.* It is likely that the next edition, which is being developed now, will reflect a quite different approach to accommodations and modifications. First, it is likely to clarify the distinctions between the two. It is also likely to no longer include a separate chapter on testing individuals with disabilities, but instead will include that discussion (along with a discussion of testing individuals of diverse linguistic backgrounds) in one chapter on fairness. It is also likely to refer to the notions of accessibility and universal design, encompassing not just individuals with disabilities or limited English, but rather that fairness and accessibility issues are relevant for all individuals.

> A document that addresses operational practices for state assessment programs (CCSSO & ATP, 2010) includes the concept of universal design, noting that it may be defined as: A set of assessment construction principles that seeks to maximize accessibility of the assessment for all students by avoiding content that may create distractions or irrelevancies for some test-takers, especially those who fall within a special population. (p. 130)

It is noteworthy that this document as well recognizes that accessibility applies to students beyond those in "special populations," although it does focus specifically on best practice statements that apply to "special populations."

Lessons Learned from Research that Applies to All Students

It is rare to find research or discussions of accommodations for students without disabilities, except when they are used as a comparison group to determine whether accommodations produce a differential boost for students with disabilities or ELLs. Yet, with the move to universally-designed/accessible assessments, it is likely that some changes previously considered to be accommodations will now be embedded in the assessments themselves, or simply be considered as good testing practices (NCEO, 2011).

Although focused on special populations, there is evidence that some "accommodations" do address needs that exist in students who are not in these groups. For example, there have been estimates that as many as 40% of all students may experience test anxiety (Cassady, 2010; Huberty, 2009). It has been suggested (Salend, 2011) that some accommodations can alleviate text anxiety. Recent studies seem to confirm this. For example, Hodgson et al. documented that test administrators find that providing a read aloud accommodations relieves student's anxiety about the testing situation. Feldman, Kim, and Elliott (2011) found that there was a relationship between self-efficacy and performance on assessments for all students, and further that accommodations packages increased the test-related self efficacy for all students, although more so for students with disabilities. And, Stowell and Bennett (2010) found that some students experienced less text anxiety when they took an assessment online rather than taking a paper-pencil test.

Despite some beginning evidence of the positive effects of accommodations for students without disabilities, there is a belief among some that it is just not reasonable to provide accommodations to students other than those legally entitled to them (Koretz, 2008). This perspective probably grows out of concerns about the logistics of providing accommodations. Others have indicated that "a person's disability status is much less important than the functional impairment that she or he may experience in trying to demonstrate proficiency on an assessment" (Kettler, 2012, p. 54).

Conclusions

The evolution that has occurred in the testing of students with disabilities and ELLs over the past two decades is truly amazing. Starting from a perspective that including these students was just too difficult, the field has moved as a result of both policy directives and advocacy to a point where the inclusion of these students is expected and expected to be appropriate for them to show their knowledge and skills. This has meant both the provision of accommodations, and the development of assessments that are more appropriate for all students.

The evolution in assessments continues as states move away from having their own state assessments to where they are part of a consortium of states that share a common assessment (Center for K–12 Assessment and Performance Management at ETS, 2012). It remains a question as to the extent to which these assessments will go as far as they might in following the principles of universal design and accessibility. Another question is the extent to which these new assessments will recognize the emerging perspective that just like accessibility, accommodations may be appropriate for a wider range of students than just those with disabilities, limited English proficiency, or both disabilities and limited English proficiency.

It will be important for researchers and assessment programs to re-evaluate whether the theoretical perspective of differential boost makes sense for groups defined by disability, or category of disability, and instead define it by student need (or functional impairment, in Kettler's, 2012, terminology). Emerging evidence suggests that students with and without disabilities do benefit from certain accommodations, just as students in both groups do not benefit from the same accommodations (see Thurlow, Moen, Lekwa, & Scullin, 2010); in other words, some students with disabilities and some students without disabilities show differential boost compared to students with and without disabilities who do not need the same accommodation. This concern about using differential boost based on group characteristics rather than based on individual student need is consistent with Kettler's (2012) conclusion:

> Researchers can improve upon the studies already conducted by moving beyond the federal distinction of disability to selecting treatment group participants based on common functional impairments. (p. 61)

Finally, it also will be important to address inconsistencies between K–12 and post-secondary educational and work settings. As students who have used accommodations move toward college and career, there should be a smooth transition. Although in many ways, postsecondary settings are more "accommodating" than are K–12 assessment environments (Gregg, 2009; Thurlow, Johnstone, & Ketterlin-Geller, 2008), entrance exams continue to place restrictions on which students may have access to accommodations, thereby possibly limiting the number of students who enter (United States Government Accountability Office, 2011). The same questions should emerge for post K–12 assessments as emerge for K–12 assessments. To what extent can we ensure that all students have access to embedded assessment features in the name of universal design, and to what extent can all students who need them have access to

accommodations so that their performance on assessments truly reflects what they know and can

do?

References

Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher, 33*(1), 1–14.

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

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

Abedi, J., Leon, S., & Kao, J. (2008b). *Examining differential item functioning in reading assessments for students with disabilities* (CSE Report 744). Los Angeles: University of California, Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, and Student Testing.

Abedi, J., Lord, C., Hofstetter, C., & Baker, E. (2000). Impact of accommodation strategies on English language learners' test performance. *Educational Measurement: Issues and Practice, 19*(3), 16–26.

Almond, P.J., Cameto, R., Johnstone, C.J., Laitusis, C., Lazarus, S., Nagle, K., Parker, C., Roach, A.T., & Sato, E. (2009). *White paper: Cognitive interview methods in reading test design and development for alternate assessments based on modified academic achievement standards (AA-MAS)*. Dover, NH: Measured Progress and Menlo Park, CA: SRI International.

Almond, P., Winter, P., Cameto, R., Russell, M., Sato, E., Clarke, J., Torres, C., Haertel, G., Dolan, R., Beddow, P., & Lazarus, S. (2010). *Technology enabled and universally designed assessment: Considering access in measuring the achievement of students with disabilities – A foundation for research.* Dover, NH: Measured Progress and Menlo Park, CA: SRI International.

Altman, J. R., Lazarus, S. L., Thurlow, M. L., Quenemoen, R. F., Cuthbert, M., & Cormier, D. C. (2008). *2007 survey of states: Activities, changes, and challenges for special education*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education (AERA/APA/NCME). (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Anderson, N.E., Jenkins, F.F., & Miller, K.E. (1996). *NAEP inclusion criteria and testing accommodations: Findings from the NAEP 1995 field test in mathematics.* Washington, DC: National Center for Education Statistics.

Bennett, R.E., Rock, D.A., & Jirle, T. (1986, February). *The psychometric characteristics of the GRE General Test for three handicapped groups* (ETS Research Report RR-86-6). Princeton, NJ: Educational Testing Service.

Bennett, R.E., Rock, D.A., & Kaplan, B.A. (1985, November). *The psychometric characteristics of the SAT for nine handicapped groups* (ETS Research Report RR-85-49). Princeton, NJ: Educational Testing Service.

Bielinski, J., Sheinker, A., & Ysseldyke, J. (2003). *Varied opinions on how to report accommodated test scores: Findings based on CTB/McGraw-Hill's framework for classifying accommodations* (Synthesis Report 49). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Bolt, S., & Roach, A.T. (2009). *Inclusive assessment and accountability: A guide to accommodations for students with diverse needs.* New York: Guilford Press.

Bolt, S.E., & Ysseldyke, J.E. (2006). Comparing DIF across math and reading/language arts tests for students receiving a read-aloud accommodation. *Applied Measurement in Education, 19*(4), 329–355.

Bowen, S., & Ferrell, K. (2003). Assessment in low-incidence disabilities: The day-to-day realities. *Rural Special Education Quarterly, 22*(4), 10–19.

Burns, E. (1998). *Testing accommodations for students with disabilities.* Springfield, IL: Charles C. Thomas.

Cahalan-Laitusis, C., Morgan, D., Bridgeman, B., Zanna, J., & Stone, E. (2007). *Examination of fatigue effects from extended time accommodations on the SAT Reasoning Test.* New York: College Board.

Cameto, R., & Nagle, K., (2010). *Where are we going and what have we learned? Research related to alternate assessments based on alternate achievement standards.* (Structured Poster Session). Presented at the annual meeting of the American Educational Research Association, Denver.

Cassady, J.C. (2010). Text anxiety: Contemporary theories and implications for learning. In J.C.

Cassady (Ed.), *Anxiety in schools: The causes, consequences, and solutions for academic anxieties* (pp. 7–26). New York: Peter Lang.

CAST. (2009). *UDL guidelines.* Retrieved from www.cast.org/publications/UDLguidelines. Center for K-12 Assessment and Performance Management at ETS. (2012). *Coming together to raise achievement: New assessments for the Common Core State Standards.* Princeton, NJ: Educational Testing Service.

Christensen, L.L., Albus, D., & Thurlow, M.L. (in press). *Accommodation policies for ELLs with disabilities.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes, IVARED.

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

Christensen, L.L., Carver, W., & VanDeZande, J., & Lazarus, S. (2011). *Accommodations manual: How to select, administer, and evaluate the use of accommodations for instruction and assessment of students with disabilities* (3rd. ed.). Washington, DC: Council of Chief State School Officers, Assessing Special Education Students State Collaborative on Assessment and Student Standards.

Christensen, L. L., Lazarus, S. S., Lail, K., Crone, M., & Thurlow, M. L. (2008). *2007 state policies on assessment participation and accommodations for students with disabilities* (Synthesis Report 69). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Christensen, L.L., Liu, K.K., & Thurlow, M.L. (2010). Professional development for teaching ELLs with disabilities. In C.J. Casteel & K.G. Ballantyne (Eds.), *Professional development in action: Improving teaching for English learners.* Washington, DC: National Clearinghouse for English language Acquisition and Language Instruction Educational Programs.

Christensen, L.L., Thurlow, M.L., & Wang, T. (2009). *Improving accommodations outcomes: Monitoring instructional and assessment accommodations for students with disabilities.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Clapper, A.T., Morse, A.B., Lazarus, S.S., Thompson, S.J., & Thurlow, M.L. (2005). 2003 state policies on assessment participation and accommodations for students with disabilities (Synthesis Report 56). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Clapper, A. T., Morse, A. B., Thompson, S. J., & Thurlow, M. L. (2005). Access assistants for state assessments: A study of state guidelines for scribes, readers, and sign language interpreters (Synthesis Report 58). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Cook, L., Eignor, D., Steinberg, J., Sawaki, Y., & Cline, F. (2009). Using factor analysis to investigate the impact of accommodations on the scores of students with disabilities on a reading comprehension assessment. *Journal of Applied Testing Technology, 10*(2).

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.

Cortiella, C. (2009). *A parent's guide to universal design for learning: Understanding the Standards-based IEP.* New York: National Center for Learning Disabilities.

Cortiella, C., & Kaloi, L. (2009). *Understanding the Americans with Disabilities Act Amendments Act and Section 504 of the Rehabilitation Act: The impact on students with LD and AD/HD* (Parent Advocacy Brief). New York: National Center for Learning Disabilities.

CCSSO (Council of Chief State School Officers) & ATP (Association of Test Publishers). (2010). *Operational best practices for statewide large-scale assessment programs.* Washington, DC: Authors.

Cox, M.L., Herner, J.G., Demczyk, M.J., & Nieberding, J.L. (2006). Provision of testing accommodations for students with disabilities on statewide tests: Statistical links with participation and discipline rates. *Remedial and Special Education, 27*(6), 346–354.

Crawford, L. (2007). *State testing accommodations: A look at their value and validity.* New York: National Center for Learning Disabilities.

DeStefano, L., Shriner, J.G., & Lloyd, C.A. (2001). Teacher decision making in participation of students with disabilities in large-scale assessment. *Exceptional Children, 68*(1), 7–22.

Dolan, R.P., Burling, K.S., Harms, M., Beck, R., Hanna, E., Jude, J., Murray, E.A., Rose, D.H., & Way, W. (2009). *Universal design for computer-based testing guidelines.* Iowa City, IA: Pearson.

Dolan, R.P., Hall, T.E., Banerjee, M., Chun, E., & Strangman, N. (2005). Applying principles of universal design to test delivery: The effect of computer-based read-aloud on test performance of high school students with learning disabilities. *Journal of Technology, Learning, and Assessment, 3*(7), 1–31.

Elliott, J., Bielinski, J., Thurlow, M., DeVito, P., & Hedlund, E. (1999). *Accommodations and the performance of all students on Rhode Island's performance assessment* (Rhode Island Report 1). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Elliott, J.L., & Thurlow, M.L. (2006). Improving test performance of students with disabilities on district and state assessments (2nd. ed.). Thousand Oaks, CA: Corwin.

Erdogan, Y. (2008). Paper-based and computer-based concept mapping: The effects on computer achievement, computer anxiety and computer attitude. *British Journal of Educational Technology, 40*(5), 821–836.

Feldman, E., Kim, J.S., & Elliott, S.N. (2011). The effects of accommodations on adolexcents' self-efficacy and test performance. *The Journal of Special Education, 45*(2), 77–88.

Finch, H., Barton, K., & Meyer, P. (2009). Differential item functioning analysis for accommodated versus non-accommodated students. *Educational Assessment, 14*(1), 38–56.

Fletcher, J.M., Francis, D.J., Boudousquie, A., Copeland, K., Young, V., Kalinowski, S., & Vaughn, S. (2006). Effects of accommodations on high-stakes testing for students with reading disabilities. *Exceptional Children, 72*(2), 136–150.

Fletcher, J.M., Francis, D.J., O'Malley, K., Copeland, K., Mehta, P., Caldwell, C.J., Kalinowski, S., Young, V., & Vaughn, S. (2009). Effects of a bundled accommodations package on high-stakes testing for middle school students with reading disabilities. *Exceptional Children, 75*(4), 447–463.

Fuchs, L.S., & Fuchs, D. (1999). Fair and unfair testing accommodations. *School Administrator, 56*(10), 24–29.

Gregg, N. (2009). *Adolescents and adults with learning disabilities and ADHD: Assessment and accommodation.* New York: Guilford.

Gregg, N. (2011). Increasing access to learning for the Adult Basic Education learner with learning disabilities: Evidence-based accommodation research. *Journal of Learning Disabilities,45*(1), 47–63.

Hodgson, J.R., Lazarus, S.S., Price, L., Altman, J.R., & Thurlow, M.L. (2012). *Test administrators' perspectives on the use of the read aloud accommodation on state tests for accountability* (Technical Report). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Hodgson, J.R., Lazarus, S.S., & Thurlow, M.L. (2011). *Professional development to improve accommodations decisions—A review of the literature*. (Synthesis Report 84). Minneapolis MN: University of Minnesota, National Center on Educational Outcomes.

Hollenbeck, K. (2002). Determining when test alterations are valid accommodations or modifications for large-scale assessment. In G. Tindal & T. Haladyna (Eds.), *Large-scale assessment programs for all students* (pp. 109–148). Mahwah, NJ: Lawrence Erlbaum.

Hollenbeck, K., Tindal, G., & Almond, P. (1998). Teachers' knowledge of accommodations as a validity issue in high-stakes testing. *The Journal of Special Education, 32,* 175–183.

Huberty, T.J. (2009). Test and performance anxiety. *Principal Leadership, 10*(1), 12–16.
Johnstone, C.J., Altman, J., & Thurlow, M.L. (2006). *A state guide to the development of universally designed assessments.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

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

Johnstone, C., Liu, K., Altman, J., & Thurlow, M. (2007). *Student think aloud reflections on comprehensible and readable assessment items: Perspectives on what does and does not make an item readable* (Technical Report 48). Minneapolis, MN: National Center on Educational Outcomes.

Johnstone, C.J., Thompson, S.J., Miller, N.A., & Thurlow, M.L. (2008). Universal design and multi-method approaches to item review. *Educational Measurement: Issues and Practice, 27*(1), 25–36.

Kamei-Hannan, C. (2008, May). Examining the accessibility of a computerized adapted test using assistive technology. *Journal of Visual Impairment and Blindness,* 261–271.

Kantrowitz, B., & Springen, K. (1997, Oct 6). Why Johnny stayed home: On test days, some schools encourage absences. *Newsweek, 130*(14), 60.

Kato, K., Moen, R., & Thurlow, M. (2009). Differentials of a state reading assessment: Item functioning, distractor functioning, and omission frequency for disability categories. *Educational Measurement: Issues and Practice, 28*(2), 28–40.

Ketterlin-Geller, L.R. (2005). Knowing what all students know: Procedures for developing universal design for assessment. *The Journal of Technology, Learning, and Assessment, 4*(2), 1–21.

Ketterlin-Geller, L.R. (2008). Testing students with special needs: A model for understanding the interaction between assessment and student characteristics in a universally designed environment. *Educational Measurement: Issues and Practice, 27*(3), 3–16.

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.

Kettler, R.J. (2012). Testing accommodations: Theory and research to inform practice. *International Journal of Disability, Development and Education, 59*(1), 53–66.

Kim, D.H., Schneider, C., & Siskind, T. (2009). Examining the underlying factor structure of a statewide science test under oral and standard administrations. *Journal of Psychoeducational Assessment, 27*(4), 323–333.

Kingston, N.M. (2009). Comparability of computer- and paper-administered multiple-choice tests for K-12 populations: A synthesis. *Applied Measurement in Education, 22*(1), 22–27.

Koenig, J.A., & Bachman, L.F. (Eds.). (2004). *Keeping score for all: The effects of inclusion and accommodation policies on large-scale educational assessments.* Washington, DC: National Academies Press.

Kopriva, R.J., Emick, J.E., Hipolito-Delgado, C.P., & Cameron, C.A. (2007). Do propoer accommodation assignments make a difference? Examining the impact of improving decision making on scores for English Language Learners. *Educational Measurement: Issues and Practice, 26*(3), 11–20.

Koretz, D. (2008). *Measuring up: What educational testing really tells us.* Cambridge, MA: Harvard University Press.

Koretz, D.M., & Barron, S.L. (1998). *The validity of gains in scores on the Kentucky Instructional Results Information System (KIRIS).* Santa Monica, CA: Rand Corporation.

Koretz, D.M., & Hamilton, L. (2000). Assessing students with disabilities in Kentucky: Inclusion, student performance, and validity. *Educational Evaluation and Policy Analysis, 22*(3), 255–272.

Laing, J., & Farmer, M. (1984). *Use of the ACT assessment by examinees with disabilities* (Research Report No. 84). Iowa City, IA: American College Testing Program.

Laitusis, C.C. (2007). Research designs and analysis for studying accommodations on assessments. In C.C. Laitusis & L.L. Cook (Eds.), *Large-scale assessment and accommodations: What works?* (pp. 67–79). Arlington, VA: Council for Exceptional Children.

Laitusis, C.C. (2010). Examining the impact of audio presentation on tests of reading comprehension. *Applied Measurement in Education, 23*, 153–167.

Laitusis, C.C., & Cook, L.L. (2007). *Large-scale assessment and accommodations: What works?* Arlington, VA: Council for Exceptional Children.

Langley, J., & Olsen, K. (2003). Training district and state personnel on accommodations: A study of state practices, challenges, and resources. Washington, DC: Council of Chief State School Officers.

Lazarus, S. S., Thurlow, M. L., Lail, K. E., Eisenbraun, K. D., & Kato, K. (2006). *2005 state policies on assessment participation and accommodations for students with disabilities* (Synthesis Report 64). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Liu, K., Albus, D., & Thurlow, M. (2006). Examining participation and performance as a basis for improving performance. *Journal of Special Education Administration. 19*(1): 34–42.

Madaus, G., Russell, M., & Higgins, J. (2009). *The paradoxes of high stakes testing: How they affect students, their parents, teachers, principals, schools, and society.* Charlotte, NC: Information Age Publishing.

Mazzeo, J., Carlson, J.E., Voekl, K.E., & Lutkus, A.D. (2000). *Increasing the participation of special needs students in NAEP* (NCES 2000-473). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

McDonnell, L.M., McLaughlin, M.J., & Morison, P. (Eds.). (1997). *Educating one & all: Students with disabilities and standards-based reform.* Washington, DC: National Academy Press.

McKevitt, B., Marquart, A., Mroch, A., Schulte, A.G., Elliott, S.N., & Kratochwill, T.R. (2000). *Understanding the effects of testing accommodations: A single case approach.* Paper presented at the annual meeting of the Council of Chief State School Officers, Snowbird, UT.

Meyen, E., Poggio, J., Seok, S., & Smith, S. (2006). Equity for students with high-incidence disabilities in statewide assessments: A technology-based solution. *Focus on Exceptional Children, 38*(7), 1–8.

NCEO. (2011). *Don't forget accommodations! Five questions to ask when moving to technology-based assessments* (NCEO Brief Number 1). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Nolet, V., & McLaughlin, M.J. (2005). *Accessing the general curriculum: Including students with disabilities in standards-based reform* (2nd. ed.)*.* Thousand Oaks, CA: Corwin Press.

Pennock-Roman, M., & Rivera, C. (2011). Mean effects of test accommodations for ELLs and non-ELLs: A meta-analysis of experimental studies.

Phillips, S.E. (1994). High-stakes testing accommodations: Validity versus disabled rights. *Applied Measurement in Education, 7*92), 93–120.

Rhode Island Department of Education. (2003). *Rhode Island assessment accommodation study: Research summary.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

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

Rogers, C.M., Christian, B., & Thurlow, M.L. (2012). *A summary of the research on the effects of test accommodations: 2009-2010.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Russell, M. (1999). Testing on computers: A follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives, 7*(20).

Russell, M. (2011). *Digital test delivery: Empowering accessible test design to increase test validity for all students*. Washington, DC: Arabella Advisors.

Russell, M., Hoffman, R., & Higgins, J. (2009a). Nimble Tools: A universally designed test delivery system. *Teaching Exceptional Children, 42*(2), 6–12.

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

Salend, S.J. (2011). Addressing test anxiety. *Teaching Exceptional Children, 44*(2), 58–68.

Shah, N. (2012, April 3). "Blind Oregon students now take online adaptive tests." *Education Week.*

Shriner, J.G., & DeStefano, L. (2003). Participation and accommodation in state assessment: The role of individualized education programs. *Exceptional Children, 69*(2), 147–161.

Shriner, J.G., Spande, G.E., & Thurlow, M.L. *State special education outcomes 2003: A report on state activities in the assessment of educational outcomes for students with disabilities.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Sireci, S.G., Scarpati, S.E., & Li, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. *Review of Educational Research, 75*(4), 457–490.

Stowell, J.R., & Bennett, D. (2010). Effects of online testing on student exam performance and test anxiety. *Journal of Educational Computing Research, 42,* 161–171.

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

Thompson, S.J., Quenemoen, R., & Thurlow, M.L. (2006). Factors to consider in the design of inclusive online assessments. In M. Hricko (Ed.), *Online assessment and measurement: Foundations and challenges* (pp. 102–117). Hershey, PA: Information Science Publishing.

Thompson, S.J., Thurlow, M.L., & Malouf, D. (2004, May). Creating better tests for everyone through universally designed assessments. *Journal of Applied Testing Technology, 10*(2). See **http:// www.testpublishers.org/atp.journal.htm**.

Thurlow, M.L., Albus, D., & Christensen, L.L. (2009). *Minnesota manual of accommodations for students with disabilities in instruction and assessment: A guide to selecting, administering, and evaluating the use of accommodations.* Roseville, MN: Minnesota Department of Education.

Thurlow, M.L., Elliott, J.L., & Ysseldyke, J.E. (1998). *Testing students with disabilities: Practical strategies for complying with district and state requirements.* Thousand Oaks, CA: Corwin Press.

Thurlow, M. L., Elliott, J.E., & Ysseldyke, J.E. (2003). Testing students with disabilities: Practical strategies for complying with district and state requirements (2nd. ed.). Thousand Oaks, CA: Corwin.

Thurlow, M.L., House, A., Boys, C., Scott, D., & Ysseldyke, J. (2000). *State participation and accommodations policies for students with disabilities: 1999 update* (Synthesis Report 33). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M., Johnstone, C., & Ketterlin-Geller, L. (2008). Universal design of assessment. In S. Burgstahler & R. Cory (Eds.), *Universal design in post-secondary education: From principles to practice* (pp. 73–81). Cambridge, MA: Harvard Education Press.

Thurlow, M., Johnstone, C., Thompson, S., & Case, B. (2008). Using universal design research and perspectives to increase the validity of scores on large-scale asessments. In R.C. Johnson & R.E. Mitchell (Eds.), *Testing deaf students in an age of accountability* (pp. 63–75). Washington, DC: Gallaudet University Press.

Thurlow, M. L., Laitusis, C. C., Dillon, D. R., Cook, L. L., Moen, R. E., Abedi, J., & O'Brien, D. G. (2009). *Accessibility principles for reading assessments.* Minneapolis, MN: National Accessible Reading Assessment Projects.

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

Thurlow, M.L., Lazarus, S.S., & Christensen, L.L. (2013). Accommodations for assessment. In B. Cook & M. Tankersley (Eds.), *Effective practices in special education* (pp. 31–1327). Iowa City: Pearson.

Thurlow, M.L., Lazarus, S., Thompson, S., & Robey, J. (2002). *2001 state policies on assessment participation and accommodations* (Synthesis Report 46). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M.L. & Liu, K.K. (2001). Can "all" really mean students with disabilities who have limited English proficiency? *Journal of Special Education Leadership, 14* (2), 63–71.

Thurlow, M. L., McGrew, K.S., Tindal, G., Thompson, S. L., Ysseldyke, J. E., & Elliott, J. L. (2000). Assessment accommodations research: Considerations for design and analysis (Technical Report 26). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M.L., Moen, R.E., Lekwa, A.J., Scullin, S.B., (2010). *Examination of a reading pen as a partial auditory accommodation for reading assessment. Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.*

Thurlow, M.L., Quenemoen, R.F., & Lazarus, S.S. (2011). Meeting the needs of special education students: Recommendations for the Race to the Top consortia and states. Washington, DC: Arabella Advisors.

Thurlow, M.L., Scott, D., & Ysseldyke, J.E. (1995). *A compilation of states' guidelines for accommodations in assessment for students with disabilities* (Synthesis Report 18). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M. L., Seyfarth, A. L., Scott, D. L., & Ysseldyke, J. E. (1997). *State assessment policies on participation and accommodations for students with disabilities: 1997 update* (Synthesis Report 29). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M. L., Shriner, J. G., & Ysseldyke, J. E. (1994). *Students with disabilities in the context of educational reform based on statewide educational assessments.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Thurlow, M. L., Ysseldyke, J. E., & Silverstein, B. (March, 1993). *Testing accommodations for students with disabilities: A review of the literature* (Synthesis Report 4). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M. L., Ysseldyke, J. E., & Silverstein, B. (1995). Testing accommodations for students with disabilities. *Remedial and Special Education, 16* (5), 260–270.

Tindal, G., & Fuchs, L. (1999). *A summary of research on test accommodations: An empirical basis for defining test accommodations*. Lexington, KY: Mid-South Regional Resource Center. (ERIC Document Reproduction Service No. ED 442 245).

United States Government Accountability Office (GAO). (2011). *Higher education and disability: Improved federal enforcement needed to better protect students' rights to testing accommodations* (GAO-12–40). Washington, DC: Author.

Willingham, W.W., Ragosta, M., Bennett, R.E., Braun, H., Rock, D.A., & Powers, D.E. (Eds.). (1988). *Testing handicapped people.* Boston: Allyn & Bacon.

Wu, Y.C., Liu, K.K., Thurlow, M.L., Lazarus, S.S., Altman, J., & Christian, E. (2012). *Characteristics of low performing special education and non-special education students on large-scale assessments* (Technical Report 60). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Ysseldyke, J. E., Algozzine, B. A., Shinn, M., & McGue, M. (1982). Similarities and differences among low achievers and students labeled learning disabled. Journal of Special Education, 16, 73–85

Ysseldyke, J. E., Algozzine, B., & Thurlow, M. L. (1992). *Critical issues in special education*. Boston: Houghton-Mifflin.

Ysseldyke, J.E., Thurlow, M.L., & Linn, D. (1999). *NGA Brief: Including students with disabilities in statewide assessments and accountability systems*. Washington, DC: National Governors' Association.

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.

Zlatos, B. (1994). Don't test, don't tell: Is "academic red-shirting" skewing the way we rank our schools? *American School Board Journal, 181*(11), 24–28.

Table 1. Number of States Allowing and Prohibiting Selected Accommodations

Accommodation	Allowed	Allowed with Restrictions	Prohibited	No Policy
Braille Edition	47	3	0	0
Read Aloud Passage	2	21	9	18
Prompt/Encourage	13	2	0	35
Calculator	10	33	4	3
Scribe	35	15	0	0
Spell Checker	8	13	16	13
Extended Time	38	8	0	4
Over Multiple Days	15	5	2	28
Individual	47	0	0	3
Small Group	48	0	0	2

Source: Christensen et al., 2011.

Table 2. Elements of Universally Designed Assessments

Element	Explanation
Inclusive Assessment Population	Tests designed for state, district, or school accountability must include every student except those in the alternate assessment, and this is reflected in assessment design and field testing procedures.
Precisely Defined Constructs	The specific constructs tested must be clearly defined so that all construct irrelevant cognitive, sensory, emotional, and physical barriers can be removed.
Accessible, Non-Biased Items	Accessibility is built into items from the beginning, and bias review procedures ensure that quality is retained in all items.
Amenable to Accommodations	The test design facilitates the use of needed accommodations (e.g., all items can be Brailled).
Simple, Clear, and Intuitive Instructions and Procedures	All instructions and procedures are simple, clear, and presented in understandable language.
Maximum Readability and Comprehensibility	A variety of readability and plain language guidelines are followed (e.g., sentence length and number of difficult words are kept to a minimum) to produce readable and comprehensible text.
Maximum Legibility	Characteristics that ensure easy decipherability are applied to text, to tables, figures, and illustrations, and to response formats.

Source: Thompson, Thurlow, & Malouf, 2004.

Table 3. Number of States' Written Policies that Allow Various Groups of Students to Use Assessment Accommodations

Group/Year	2002–2003	2004–2005	2006–2007	2008–2009
Students with Disabilities	50	50	50	50
504 Students	33	41	45	42
ELLs[a]	NA	NA	45	42
ELL/IEP	13	25	45	48
All Students[b]	15	9	14	15

Sources: Christensen, Braam, Scullin, & Thurlow, 2011; Christensen, Lazarus, Lail, Crone, & Thurlow, 2008; Clapper, Morse, Lazarus, Thompson, & Thurlow (2005); Lazarus, Thurlow, Lail, Eisenbraun, & Kato, 2006.

[a] In analyses of early accommodations guidelines, references to ELLs were not documented.

[b] In early years, "all students" included students receiving Title I services. In addition, across years, states often made a distinction between accommodations allowed for all students without qualifications, and those with qualifications; qualifications included, for example, that there must be a documented need.

Table 4. Student Characteristics, Needs, and Possible Assessment Accommodations

Student Characteristic	Need	Possible Accommodation
Weak manual dexterity; difficulty with pencil; difficulty typing on standard keyboard	Assistance with holding materials, using pencil, typing	Express response via scribe or pointing Use speech to text technology
Poor decoding skills	Assistance with reading when the skills measured is not decoding skills	Human reader Audio tape or CD Text to speech technology
Short attention span; easily distracted	Assistance with focusing on assessment or key parts of assessment	Highlighting of key words Masking or template to mark place Location near teacher of proctor Human reader or text to speech technology

Sources: Thurlow, Albus, & Christensen, 2009.