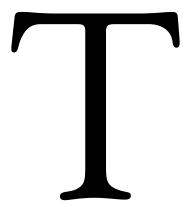# ASSESSMENT ACCOMMODATIONS ON TESTS OF ACADEMIC ACHIEVEMENT FOR STUDENTS WHO ARE DEAF OR HARD OF HEARING: A QUALITATIVE META-ANALYSIS OF THE RESEARCH LITERATURE

STEPHANIE CAWTHON AND RACHEL LEPPO

BOTH AUTHORS ARE AFFILIATED WITH THE DEPARTMENT OF EDUCATIONAL PSYCHOLOGY, UNIVERSITY OF TEXAS, AUSTIN. CAWTHON IS AN ASSOCIATE PROFESSOR AND LEPPO IS A GRADUATE RESEARCH ASSISTANT.

THE AUTHORS conducted a qualitative meta-analysis of the research on assessment accommodations for students who are deaf or hard of hearing. There were 16 identified studies that analyzed the impact of factors related to student performance on academic assessments across different educational settings, content areas, and types of assessment accommodations. The meta-analysis found that the results of analyses of group effects of accommodated versus unaccommodated test formats are often not significant, test-level factors exist that can affect how students perceive the assessments, and differences exist in how test items function across different conditions. Student-level factors, including educational context and academic proficiency, influence accommodations' role in assessment processes. The results of this analysis highlight the complexity of and intersections between student-level factors, test-level factors, and larger policy contexts. Findings are discussed within the context of larger changes in academic assessment, including computer-based administration and high-stakes testing.

Keywords: accommodations, assessment, student factors

Educational testing in U.S. public education continues to shift toward two goals: (a) standardized measurement of student outcomes and (b) inclusive assessment participation policies. The first goal allows for comparisons of student progress across different settings. The scope of comparisons varies widely: Individual schools in a state can be compared, or, per one of the goals of the upcoming Common Core Standards Assessments, individual states across the nation (Christensen, Lazarus, Crone, & Thurlow, 2008; Common Core State Standards Initiative, 2010; Rigney, Wiley, & Kopriva, 2008; Thurlow, Lazarus, Thompson, & Robey, 2002). At the same time, inclusive assessment participation policies require that all students be provided with an opportunity to demonstrate their knowledge and skills on these tests. Expansion of the set of participants in an assessment can be challenging to implement in a fair and valid way when tests were not

363

originally designed to be accessible to all students, particularly students with disabilities or those who are English Language Learners (Cawthon, Ho, Patel, Potvin & Trundt, 2009; Kopriva, Emick, Hipolito-Delgado, & Cameron, 2007; Shaftel, Belton-Kocher, Glassnap, & Poggio, 2006; Sireci, Li, & Scarpati, 2003).

In the United States, both state-level and federal education policies (Individuals With Disabilities Education Improvement Act, 2004; Americans With Disabilities Act, 1990) govern how students with disabilities participate in large-scale assessments. Within the context of accountability reforms such as NCLB, tests are high-stakes activities because the resultant scores drive decisions about student promotion and school accountability report cards, and in some states, teacher compensation. The impact of decisions that include consideration of standardized assessments thus reaches across all levels of the education system, providing a significant impetus to facilitate inclusive participation of diverse student populations in a way that yields results that are both meaningful and valid (Elliott & Roach, 2007; Fairbain & Fox, 2009; Kopriva, 2008; Liu & Anderson, 2008; Shaftel, Yang, Glassnap, & Poggio, 2006).

Assessment accommodations are one strategy frequently used to expand the inclusivity of standardized assessments for students with disabilities (Lazarus, Thurlow, Lail, Eisenbraun, & Kato, 2006). Common accommodations allowed on state standardized assessments include extended time, a separate room for administration of the test, having test items read aloud (for tests that are not in a language in which a student reads proficiently), and having test instructions read aloud before the student begins the assessment (Christensen, Braam, Scullin, & Thurlow, 2011). Students may receive

test accommodations singly, or together as a package. In contrast with test *modifications*, in which changes to the test format or content may alter the construct being measured, assessment *accommodations* are meant to increase access to the test content while allowing for the score to be interpreted in the same manner as that for a test taken without an accommodation (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). Maintaining the construct of the test item, and thus the validity of the test score, is critical when scores are being used for high-stakes decisions such as those derived from NCLB.

While state assessment policies seek to offer clear guidance on the use of accommodations in high-stakes assessments, the body of empirical research literature examining the actual effects of test accommodations on student test scores offers less clarity as to accommodations' potential impact on test score validity (Abedi, Hofstetter, & Lord, 2004; Bolt & Thurlow, 2004). While there are examples of test accommodations having the desired effect of providing access without changing test difficulty (Fletcher et al., 2006; Schulte, Elliott, & Kratochwill, 2001), there are also examples of test accommodations that have the potential to change the construct of the test item (Crawford & Tindal, 2004; Fletcher et al., 2006), and that sometimes even make the test item more difficult for the student using the accommodation (Sireci, Scarpati, & Li, 2005).

Comprehensive reviews of research findings on the effects of test accommodations on student scores are available for students with disabilities as a whole (Bolt & Thurlow, 2004; Towles-Reeves, Kleinert, & Muhomba, 2009; Zenisky & Sireci, 2007), and for English

Language Learners (Kopriva, 2008, pp. 81–100; Shaftel et al., 2005), as well as in disaggregated form, by type of accommodation (e.g., extended time, Elliott & Marquart, 2004; read-aloud, Meloy, Deville, & Frisbie, 2002). Common to all these studies is an appreciation of the importance of knowing the content area of the test as well as the student's proficiency in that content area when efforts are being made to understand the potential effect of an accommodation in making a test more accessible for the student (e.g., Bielinski, Ysseldyke, Bolt, J. Friedebach, & M. Friedebach, 2001; Bolt & Thurlow, 2006; Mandinach, Bridgeman, Cahalan-Laitusis, & Tripani, 2005). The reviews cited above provide excellent discussions of the nuances of the findings, the places where research findings can reasonably be translated into assessment policy, and the challenge of fitting the accommodations to the test content area as well as the student's individual needs.

For students in high-incidence disability populations, such as students who have a learning disability or attention deficit hyperactivity disorder, many of the findings in the extant reviews offer sufficient understanding of how the research literature can inform practice (Cawthon, Kaye, Lockhart, & Beretvas, 2012 ; Keiffer, Lesaux, Rivera, & Francis, 2009). However, in research on "students with disabilities" as a whole, students with low-incidence disabilities are aggregated into a larger group, which reduces the capacity to identify how an accommodation interacts with their characteristics, specifically. Diffusing findings across diverse student groups goes against one of the findings of the field: that an effective accommodation, meaning one that facilitates access to test content yet maintains the construct of the test, requires consideration of the characteristics of the student, the test,

and the accommodation (Cawthon, 2009). There are studies that specifically focus on a subgroup of students with disabilities, such as those with more severe cognitive disabilities, allowing for the kind of discussion that can meaningfully translate into assessment practice guidelines for these populations (e.g., Baker, 2010; Elliott & Roach, 2007; Zalta & Pullin, 2004). If research findings on accommodations are to be meaningfully interpreted, accommodations' effects within individual disability groups need to be considered in their own right (Sireci et al., 2005).

The present article focuses on the research literature on the effects of assessment accommodations used with students who are deaf or hard of hearing, a population that makes up less than 1% of students who receive services under the Individuals With Disabilities Education Act (see http://www.ideadata.org/partbchildcount.asp, a website of the Data Accountability Center). Students who are deaf or hard of hearing come to educational assessment with a unique but critically important set of personal characteristics and learning experiences (Cawthon, 2007). Unfortunately, much of the accommodations research literature on these students does not take into account the demographics, such as degree of hearing loss or language use, in the study designs (Cawthon, 2007; Cawthon & Online Research Lab, 2006). To continue the argument from above, although it may be tempting to group the members of this low-incidence student population together, deaf and hard of hearing students are diverse in their linguistic, communication, and cultural identification (P. E. Spencer & Marschark, 2010). There is no "typical" deaf or hard of hearing student demographic (Mitchell, 2004): Some students come to primary education fluent in American Sign Lan-

guage (ASL), others use spoken language, others a combination of visual and spoken modalities (Emmory, Petrich, & Gollan, 2013; Singleton, Morgan, DiGello, Wiles, & Rivers, 2004; Wilbur, 2000). Some students use an assistive hearing device or cochlear implant, whereas others do not use amplification devices at all (Kent & Smith, 2006; L. J. Spencer, Tomblin, & Gantz, 2012). Some identify as culturally Deaf (sometimes along with other ethnic and cultural identities) or physically deaf, or as a person who is hard of hearing (see, e.g., McIlroy & Storbeck, 2011; Nikolaraizi & Hadjikakou, 2006; Schick et al., 2013; Smiler & McKee, 2007).

The above characteristics are interrelated, and are a result of individual experiences, family history and resources, and cultural influences. Students' characteristics are tied not only to their family background, but also to where they obtain their K–12 education. Educational assessments are largely conducted within the context of school settings; deaf education has a long history that features strong ties to community resources, the importance of school identity, and a movement toward inclusive and local education settings and away from residential schools (Mitchell, 2006). The educational experiences of deaf and hard of hearing students thus reflect the great diversity of personal and community contexts in which these students live (P. E. Spencer & Marschark, 2010).

This diversity of demographics and educational experience has direct implications for how students who are deaf or hard of hearing participate in large-scale assessment systems (Cawthon, 2007, 2009), as well as curriculum-based and psychological assessments (Braden & Hannah, 1998; Luckner & Bowen, 2006; Soukup & Feinstein, 2007). Because most tests of

student academic achievement are in written English, administered either on paper or via a computer system, reading skills and literacy development are an important part of accessing the content of the assessment. English literacy skills, particularly academic literacy, vary widely in students who are deaf or hard of hearing (Qi & Mitchell, 2012). English literacy levels are an access concern for students who are reading significantly below grade level, which is the case for many students who are deaf or hard of hearing, particularly if they did not have access to a full language model during early childhood (Luetke-Stahlman & Nielsen, 2003). (For in-depth discussions and different perspectives on the linguistic determinants of English literacy development in deaf and hard of hearing students, see Czubek, 2006; Easterbrooks, Lederberg, & Connor, 2010; Kyle & Harris, 2011; Mayer, 2007; Mayer & Akamatsu, 1999; Myers et al., 2010; Padden & Ramsey, 1993; Paul, 2006; Strong & Prinz, 1997; Wilbur, 2000; Williams, 2012). Beyond reading as an access tool for test content, there are also some concerns about deaf and hard of hearing students' access to classroom learning opportunities (Moores, 2010; Vetter, Löhle, Bengel, & Burger, 2010). For example, students who use an interpreter as a classroom accommodation may not have the same level of direct instruction as students who receive direct instruction from the teacher (Cawthon, 2001; Marschark, Sapere, Convertino, & Pelz, 2008). Furthermore, students may not have access to the incidental learning that happens in a classroom between students or between the teacher and other students; this can reduce the knowledge base they can draw from at the time of the assessment.

Prevalent assessment accommodations used by students who are deaf or hard of hearing include some of the

365

same accommodations used by other students with disabilities, such as extended time, a dictionary or glossary, or a separate setting (Cawthon & Online Research Lab, 2006, 2007). Some of these accommodations function to provide support for students whose literacy level may not be at grade level and/or who may not have had full access to the content of the assessment (Cawthon, 2009). However, there are additional accommodations that are more specific to language characteristics of students who are deaf or hard of hearing, including (a) having an interpreter translate test directions, reading passages, or test items, either using sign language, a signed system, or a read-aloud approach; and (b) allowing the students to respond using sign language and having their responses recorded by a scribe who back-translates those responses into English. When an interpreter or other sign language–based accommodations are used, there are potentially significant implications for the role the resultant scores play in high-stakes decisions. Many states do not allow these kinds of scores to be aggregated with unaccommodated (or other accommodated) scores, thus reducing their viability as facilitators of inclusive assessment practices (Clapper, Morse, Lazarus, Thompson, & Thurlow, 2005). Previous studies (Cawthon, 2007, 2009, 2011) illustrate how the participation of deaf and hard of hearing students in standardized assessments in which these types of accommodations are used has potentially negative implications at both the student and school levels. At minimum, the test score is less likely to represent what the individual student has learned, and it is quite possible that decisions about the student's trajectory, teacher effectiveness, and school status may be not based on sound

information (Cawthon, Leppo, Carr, & Kopriva, 2013).

Although there is a growing body of literature on the effects of accommodations for students who are deaf or hard of hearing, there has not been a systematic analysis of results from across the field. The purpose of the present study was to conduct a qualitative meta-analysis of the research literature and to make recommendations for future research and practice. Because of the potential interaction between test, student, and accommodations factors, it was critical to examine the literature from the perspective of both student and test factors in order to understand how they may contribute to differential outcomes of test accommodations on the resultant scores (Cawthon, 2009). Three research questions guided this analysis:

1. What do research findings in the extant literature say are the effects of test accommodations for students who are deaf or hard of hearing?
2. What test-level factors are included in the extant literature on the effects of accommodations on test scores for students who are deaf or hard of hearing?
3. What student-level factors are included in the extant literature on the effects of accommodations on test scores for students who are deaf or hard of hearing?

Method
In compiling the literature review for the present article, we searched for published works that focused specifically on empirical investigations of accommodations on tests of student achievement and outcomes for students who are deaf or hard of hearing. We searched four research databases: PsycInfo, ERIC, Dissertation

Abstracts International, and Educational Abstracts. These databases cover a wide range of both education and psychology journals and other publications that are likely to have articles in this area. We searched these databases because they house two primary sources of empirical reports: scholarly journals and assessment research organizations such as the federally funded education research labs and university research centers. All research submitted to the journals undergoes a peer-review process that helps ensure study quality. We did not put a date range limit on the search; the oldest paper that met our criteria was published in 1985, and the majority had been published since 2000.

Our search terms included *accommodations*, *test*,* *assess*,* *modification(s)*, *D/deaf*, *hard of hearing*, *hard-of-hearing*, *hearing loss*, *hearing impaired*, and *deaf and hard of hearing* (or *hard-of-hearing*). The asterisk (*) is used in many search engines as a completion term to capture variations on the word. For example, *test** will return articles with *test*, *testing*, *tests*, etc. This increased the coverage of the search term to all variations that fit the word stem. Although there are technical differences between the two terms, for the purpose of this review we shall use the terms *tests* and *assessments* interchangeably. Although the focus was on assessment accommodations, there were also several articles that looked at instructional accommodation formats and the potential impact of different methods of presenting information on student test scores. We included these when the study design allowed for implications to be reasonably drawn for an assessment accommodation.

Our inclusion criteria for the analysis for the present study included the following:

- articles with experimental, quasi-experimental, or correlational designs as at least part of the methodology
- research that included students who were deaf or hard of hearing or were otherwise indicated as having a hearing loss or impairment
- studies with statistically significant and/or insignificant findings
- educational settings ranging from K–12 through postsecondary
- discussion of accommodations used during assessment, with measures of knowledge gained or information retained
- dissertation studies and published manuscripts
- studies that identified student-level variables and group membership

The studies in our literature review focused on a range of different types of assessments, including intelligence tests, instructor- or researcher-developed assessments, state standardized tests, and nationally administered assessments. The type of assessment used in each study is relevant to this discussion because it helps to provide perspective on the purpose of the test and how an accommodation (or an accommodation package) might influence test outcomes.

Exclusion criteria eliminated studies

- with instructional accommodations only, without measurement of student knowledge
- with samples of students with disabilities as a whole that did not disaggregate findings for individuals who were deaf or hard of hearing
- that considered accommodations outside an academic setting (e.g., in an employment or mental health setting)

- with participants younger than school age
- that were published in a language other than English
- that were done for a thesis and were not otherwise published in a peer-reviewed journal

Policy analyses and articles that were theoretically based were also excluded, as well as manuscripts specific to alternate assessments, specific groups of other students with disabilities, or work with English Language Learners.

Summaries of each of the included articles are provided in Table 1. The articles are organized chronologically, from the earliest to the most recent. Each summary consists of a brief overview that includes a description of the sample, measures, potential confounds, accommodations used, covariates, effects, and outcome variables. Studies were reviewed separately by both authors; in its final form, the table reflects joint decisions made during the discussion process.

Results

The three research questions (What are the effects of accommodations? What are test-level factors? What are student-level factors?) provided an overall context for the present study. The categories and themes for discussion within the sections of the article were generated as part of our review process, informed by both the research literature and issues in application to practice. In our analysis, we found that the overall effects of accommodations were mixed—a finding consistent with research reviews of accommodations with other student groups—and that the results depended greatly on the type of test and the type of accommodations used. The types of accommodations used in the reviewed studies reflect an overriding priority of providing access to English

print for students who are deaf or hard of hearing. This priority is related to the educational context of the participants; many of the students included in these studies received classroom instruction in a signed language, either via an interpreter in an inclusive setting or directly from the instructor in an education setting focused on students who are deaf or hard of hearing. However, despite the importance of matching accommodations to individual student characteristics and needs, few studies disaggregated findings beyond broad categories such as "deaf" or "DHH" or "hearing." As a result, student-level characteristics were sometimes less salient in the discussion of findings, and it was often more difficult to draw conclusions regarding the fit between students, the assessment, and the accommodations. This is a significant limitation of the research literature and raises questions as to the applicability of these findings to practice.

ASL as an Accommodation

Although students who are deaf or hard of hearing use a range of accommodations, differences between ASL (or signed) and English text presentations of information are the priority focus in many of the reviewed research studies. Translation of test items into another language raises questions about the validity of the construct being measured by the assessment. Johnson, Kimball, and Brown (2001) looked at issues surrounding interpreter quality and the student experience of using an interpreter during the assessment process. The study was a comparison of ASL and Signing Exact English (SEE II), a visual signed system with strong parallels to English syntax and phonology that is sometimes used in classroom settings. In this study of student performance on a statewide math assessment, Johnson

Table 1

Research on Accommodations and Students Who Are Deaf or Hard of Hearing

Authors (date)	Sample	Test type	Accommodation	Outcome measure
Mowl (1985)	32 hearing impaired middle school students	Statewide learning and literacy skills	Item modifications to control syntax and vocabulary	Math and reading scores
Sullivan & Schulte (1992)	368 DHH students	Intelligence test	ASL interpreter vs. signing of test items	Factor structure; scores on test and subscales
Maihoff et al. (2000)	20 DHH students	Standardized math achievement test	Signed vs. written responses to test items	Test scores
Johnson et al. (2001)	12 classrooms of students	Statewide math test	Signing Exact English (SEE II)	Test scores
Marschark et al. (2005)	College students Experiment 1: 187 DHH, rest hearing Experiment 2: 20 interpreters Experiment 3: 20 DHH, 10 hearing	Instructor-developed science test	Video vs. live interpreting of material	Test scores
Marschark et al. (2006)	College students: 105 DHH, 22 hearing	Instructor-developed math test	ASL presentation of lectures; interpreter experience and familiarity	Test scores
Ansell & Pagliaro (2006)	Early elementary: 59 Deaf students in schools for the deaf	Researcher-developed word problems	ASL interpretation of items	Problem-solving strategies
Wolf (2007)	Elementary and middle school mainstreamed settings: 62 DHH students	Stanford Achievement Test: Reading and Language (Mitchell, Qi, & Traxler, 2007)	Extended time, teacher clarification, re-reading of directions, scheduling, simplified language in directions, signed or read written directions	Test scores
Anderson-Inman et al. (2008)	Middle and high school: 9 DHH students	Instructor-developed science test	Captions vs. expanded captions	Gain scores from pretest to posttest
Stinson et al. (2009)	High school: 48 DHH students College: 48 DHH students	Instructor-developed history test	Lecture with speech-to-text vs. interpreting/note-taking services	Student retention of test material
Marschark et al. (2009)	College students Experiment 1: 20 DHH, 29 hearing Experiment 2: 23 DHH, 24 hearing	Instructor-developed science test, both multiple choice and sentence completion	Print vs. ASL/English presentation of material	Test scores
Russell et al. (2009)	98 Middle and High school DHH students	National Assessment of Educational Progress	Human vs. avatar signer	Test scores
Steinberg et al. (2009)	Fourth- and eighth-grade students: 1,000 DHH vs. more than 1 million with other disabilities or English Language Learners	State standardized math assessments	Accommodated vs. not accommodated	Item difficulty
Hoffman & Wang (2010)	First-grade students: 2 DHH	Modified test item	ASL graphics	Parent and teacher reports of student engagement
Cawthon, Wurtz, & Online Research Lab (2010)	414 K–12 teachers and administrators of DHH students	State or district standardized assessments	Extended time, test directions interpreted, test items interpreted, test items read aloud	Participant reports of accommodations use
Cawthon, Winton, et al. (2011)	Fifth- through eighth-grade students: 64 DHH	Math and reading items from state test	DVD of ASL presentation with booklet vs. paper and pencil alone	Test scores

et al. found that there were translation issues, specifically around items that had visual components, as well as open-ended questions that did not have specific multiple-choice responses. These findings emphasize the importance of item type and content when decisions are being made about the feasibility of ASL as an accommodation for high-stakes assessments.

Even with these cautions about difficulty in language translation, findings largely point to a minimal impact of ASL as an accommodation on resultant test scores. Maihoff et al. (2000) had inconclusive findings because their sample had only 20 students; also, there was not significant power to detect differences between the two test samples, and the samples were not measured with a similar assessment. Cawthon, Winton, Garberoglio, and Gobble (2011) also looked at the effects of ASL versus written text on reading and math test items taken from a state assessment (but not administered as part of the larger annual assessment). This study also did not find a significant effect of assessment format on student test scores. Interestingly, two additional factors, exposure to ASL in the classroom and student proficiency in reading, were both significant positive predictors of student test scores (overall on both ASL and written formats).

Marschark et al. (2009) looked at the effects of print versus "in the air" presentation of passages on the test performance of deaf and hard of hearing students and hearing students on a classroom science unit assessment. What makes this study unique is that it included both types of students and focused on the effects of the "in the air" format versus a printed English version of the materials. Deaf and hard of hearing students had ASL for their "in the air" format, and hearing students had auditory presentations of the passages. Across the board, all students had higher test scores on the print version of presentation than on the "in the air" format. This finding emphasized the strength of the print version of the materials over auditory or visual learning of the test content. In a follow-up experiment, Marschark et al. looked at student output on a writing passage assessment with "in the air" versus written response conditions for both student groups. Deaf and hard of hearing students scored higher on this assessment when they wrote out their responses in written format than when they responded "in the air" using ASL. Taken together, these studies suggest that ASL as an accommodation may not play an instrumental role in increasing access to written assessments for college-level students who are deaf or hard of hearing.

Computer-based assessments increase the options that may be available for video-based or digitally generated accommodations. Russell, Kavanaugh, Masters, Higgins, and Hoffmann (2009) investigated the effects of an avatar signer versus a video-recorded human ASL interpreter on student performance on the National Assessment of Educational Progress math assessment. Test scores from the NAEP are an alternative to state assessments as a tool for understanding the effects of accommodations on a nationally recognized assessment. Students in this study were in grades 8–12. Although there was no paper-and-pencil comparison format score for this assessment, students did have access to the print version of the item during the assessment. Students performed equally well on the NAEP with the avatar and video-recorded signer, a finding that suggests that the ASL presentation did not vary depending on its source. However, it is unknown if the students scored higher or lower using the ASL accommodation than they would have using only the English print version of the test.

Finally, one study used an intelligence test in its investigation of the effects of assessment accommodations on resultant scores. Sullivan and Schulte (1992) examined the difference between ASL interpretation and signing by the test administrator. The focus of this study was on the factor structure of the Wechsler Intelligence Scale for Children–Revised intelligence test and subscales (Vocabulary, Information, Comprehension, Similarities, Arithmetic, Digit Span, Object Assembly, Block Design, Picture Completion, Picture Arrangement, Coding). This study is of interest both in that it analyzed variables at the test, student, and family levels on overall outcomes and in that it looked at variations in how a visual representation of the items might affect the scores. For example, children with parents who were deaf or hard of hearing had higher mean scores on all scales. As far as the effect of the different types of administrations (ASL interpreter vs. signing administrator), there were no differences in factor structure or student profiles. There was also not an analysis of the interaction of student-level and family-level factors with the type of administration. This finding is consistent with that of Russell et al. (2009) that the type of visual modality did not appear to change the validity of the test score.

Participant Developmental Stage

Given the significant changes in language, literacy, and content-area knowledge as students move from elementary through postsecondary experiences, it is possible that the effect of accommodations may interact with the age or grade of participants. Most studies in our literature review examined student performance within a similar

age range, either in elementary or post-secondary settings. One exception, Stinson, Elliot, Kelly, and Liu (2009), investigated how speech-to-text accommodations versus interpreted lectures affected performance on multiple-choice and sentence completion tests. In their analysis across two different age groups, high school versus college students, Stinson et al. found that material retention was consistently greater for both groups on the multiple-choice test. However, for the type of accommodations used during the lecture, high school students did better with speech-to-text than with an interpreter, whereas college students did equally well under both conditions. The findings of this study suggest that the older students may not have relied on the accommodation for access in the same way as the high school students, perhaps indicating a greater flexibility in using input from ASL and written English conditions.

The analysis of test item format (multiple choice vs. sentence completion) in Stinson et al. (2009) is in line with related research on the effects of linguistic complexity on assessment scores for students who are English Language Learners (Abedi, 2006; Cawthon, 2011, Cawthon, Highley, & Leppo, 2011) or students with other disabilities (Cawthon, Beretvas, et al., 2012). Multiple-choice items instead of sentence completion tasks are not an accommodation per se, and are used on many standardized assessments. However, multiple-choice items have a lighter linguistic load and, at times, less cognitive demand than sentence completion items, which perhaps makes them easier for students who may have more emergent levels of English literacy than those who are fully proficient (and at grade level; Abedi, 2011). It may be the case that the conditions in the study by Stinson et al. (multiple choice vs. sentence completion) were

already measuring different cognitive tasks, and that the assumption that the constructs being measured were the same across assessment formats may mask or alter how the effects of the two accommodations formats can be interpreted.

Ansell and Pagliaro (2006) addressed issues of varying item difficulty and variation in student ASL proficiency on a math assessment. The tests were administered in ASL, and the investigation included measurement of the students' age and ASL skill development as factors in the analysis. The outcome of interest in this study was the type of problem-solving strategies students used on items with different levels of difficulty. This is a unique study in that the focus was as much on the way students answered the items as on whether or not the items were solved correctly. The unit test used in the study was a useful tool for this analysis because it afforded the opportunity to look closely at those different types of problem-solving strategies in a way that a standardized assessment could not. In this study, higher student ASL skill levels were related to the use of more viable problem-solving strategies on the math test items. While results were not controlled for by other developmental factors such as literacy level, and a comparable test format in written English was not used, this focused attention on the impact of ASL language level on test outcomes is a unique and valuable contribution to the field.

Single Versus Package
The majority of the studies in the present literature review looked at one accommodation at a time, either within the context of a larger investigation of the assessment process or as an independent variable compared with another single accommodation. However, in reality, most students in primary and secondary education receive

more than one assessment accommodation, or a *package* of accommodations. The studies that did look at more than one accommodation at a time varied in whether these were deliberately chosen or arose naturally from the sampling process. For example, Steinberg, Cline, Ling, Cook, and Tongatta (2009) investigated item difficulty for deaf and hard of hearing students with and without accommodations. Accommodations were not parsed out individually, but grouped as a student having one or more accommodations versus a student having none. The results of the study showed relatively few items with greater difficulty for students with accommodations than for those without, or vice versa, across the items and the grades (fourth and eighth grades). This finding suggests that the accommodation packages did not change the difficulty of the item, and thus that inferences about the test score are still valid. Such an "in situ" study is perhaps a more authentic approach to understanding the impact of an accommodated assessment than an experimental design, in practice, because it allows for accommodations to have been selected for students based on their characteristics and access needs.

Wolf (2007) also considered the effects of accommodations as a package instead of conducting an experimental investigation of a single accommodation compared with another (or none at all). In Wolf's study, students used a range of accommodations, including extended time, teacher clarification of the question, re-reading the directions, flexible scheduling, simplified language in the directions, or directions presented in a signed language or read to the student. It is important to note that all accommodations used were deemed **not** to potentially violate the construct of the item. In addition to accommodation use, both individually and as a pack-

age, Wolf included information about the degree of hearing loss as a mediating factor in this analysis, but did not include measures of language proficiency in the study. First, in the analysis of use, participants were likely to employ similar accommodations together (e.g., accommodations for directions). Second, in analysis of the impact of accommodations on test performance, the results depended on the content of the assessment. One particularly interesting finding was that the level of hearing loss and use of accommodations did not interact in their effect on test scores. In other words, there was not a differential effect of accommodations by potentially different levels of access to spoken English.

Availability Versus Actual Use

A challenge when evaluating the impact of different test accommodations is to know whether or not a student is using the accommodation as intended, and, subsequently, to make an accurate interpretation of its potential effect on a test score. In a study by Anderson-Inman, Terrazas, and Slabin (2008), middle school students took a science test with either standard captions or "expanded" captions. The "expanded" captions linked test content to additional contextual materials, such as written definitions, pictures, or concept maps. There were no reported differences in gain scores on the unit tests between the two conditions (although the sample of nine students, split between two groups, was perhaps too small to capture meaningful differences with the use of a group design). However, students did show a preference for the expanded captions, a finding that represented a common theme throughout the present literature review: that students often preferred one accommodation over another even when their performance

did not indicate differential impact on test scores (e.g., Russell et al., 2009). One of the considerations in the study by Anderson-Inman et al. was the fact that merely presenting the accommodation did not mean that the participants necessarily accessed it.

A similar question about accommodations use arose in a study by Cawthon, Winton, et al. (2011). In this study, students were provided two sets of items matched on content area and difficulty, one set provided in ASL via video and one set in print. The ASL version also included the text of the test item, so it was not the sole means of gaining access to the content of the question. The study did not include measures of eye gaze so as to ascertain the actual use of the ASL component of the video (see Marschark et al., 2005, on inclusion of eye gaze in study analysis). Students had the printed booklets with the test items, and if they felt that the ASL component was unnecessary, they could have read all of the items in both the ASL and English print formats. In a sense, this is a realistic version of what test taking with an accommodation is like for students in the classroom setting. Unless the entire assessment is administered without a booklet or test items in written format, the student is likely to have the option of reading the item in addition to or instead of using an ASL accommodation. Other than measures of student preference, student use is not controlled for in the empirical studies of accommodations without strict manipulation of the testing experience.

Discussion

Each of the reviewed articles has contributed to the growing knowledge base surrounding assessment accommodations for students who are deaf or hard of hearing. While the field is still in its infancy, even after nearly two

decades of research there are some important implications from the present review as educational assessment in the United States continues its shift toward standardization and the inclusion of all students.

Student-Level Factors

The first implication of the findings of the present review is that, in many cases, student-level factors made just as important a contribution to the research outcomes as variations in accommodations. At times, a student-level factor was as simple as whether or not the student chose (or was able) to use the target accommodation during the learning process as well as during the assessment. In other cases, a student's specific individual characteristics, including age, receptive ASL skills, academic proficiency in the test content area, and facility with written English, had a significant effect on test outcomes. Yet the extent to which studies analyzed how student characteristics varied across accommodations conditions was limited. Few studies included more than one variable in the analysis, and very little information related to etiology, degree of hearing loss, and experience with language and literacy was included in most designs. Although few studies looked at potential interaction effects of these student-level factors with accommodations use, these characteristics are an intrinsic part of how education professionals make decisions about what accommodations will be appropriate for students to use. It would appear that the examination of accommodations use by deaf and hard of hearing students now must move beyond research on changes in test scores and into examination of the mechanisms that may underlie some of the potential interactions between accommodations, test formats, and student characteristics.

371

Primary Versus Supporting Accommodations

While there are primary accommodations that are directly linked to creating access to the assessment content, there are times when the effectiveness of a target accommodation is *facilitated* by another, less emphasized accommodation within a classroom context. During class periods, for example, there are times when the instructional style makes it difficult to take notes and watch an interpreter at the same time. In this situation, a note taker may be seen as a supplementary accommodation. However, the effectiveness of having an interpreter may be somewhat contingent on the availability of a note taker who is able to provide information in an accessible format. None of the studies we reviewed specifically looked at targeted packages of accommodations, with primary and facilitated accommodations, other than to note that many students use packages (Steinberg et al., 2009) and that students who are deaf or hard of hearing are likely to use multiple types of accommodations (Wolf, 2007). Targeted research on how packages support either the instructional or the assessment process, viewed holistically and with consideration of how the package of accommodations function together, would help illustrate some potential best practices in this area.

Computer-Based Assessment

The study by Russell et al. (2009) examining avatar signers versus videotaped human signers embedded in a computer-based assessment administration represents an emerging and potentially significant area of research. A computer-based test format meets many goals in a large-scale, inclusive assessment process. First, it is efficient to implement and, once set up, holds the potential for quick, large-scale administration of tests for many students within a relatively short time. Second, it allows for tailored assignment of accommodations at the student level, and perhaps at the item level. If, as indicated above, math items that had a high visual component presented challenges to maintaining the construct being measured when presented in ASL, it would be possible simply to omit the signing component of that test item. Finally, computer-based assessment leverages the potential of computer-*adapted* assessment, such that students receive items that match their skill level as they make their way through the test. This prevents students from receiving too many items that are far above their proficiency level—which can increase frustration and reduce motivation to complete the assessment—and allows students to more efficiently demonstrate their knowledge (engagement in the assessment as facilitated by ASL graphics, discussed by Hoffman and Wang, 2010).

The current literature can inform choices about the format of a signed accommodation in a computer-based assessment. Both the Russell et al. (2009) and Marschark et al. (2005) studies show that different versions of an ASL presentation do not affect student performance. However, the findings of Russell et al. did show that students preferred the latter format to the former. Student preferences and familiarity with interpreters were also a part of the study by Marschark, Convertino, Sapere, and Seewagen (2006), but were found not to be a significant predictor of student performance. If the focus of discussion surrounding accommodations in a computer-based assessment is on differential effects of a video, live, or avatar signer, then the research literature will indicate that the specific format of the signing accommodation is not the key issue in achieving a valid test score. If the focus of the discussion is on the learning process, using computer-based signed accommodations during the classroom experience, there may be other variables and factors that will be important to include in empirical studies before drawing conclusions about the relative effectiveness of different signing options in computer-based learning formats.

Fairness

The validity of the interpretation of an accommodated test score is very important in the setting of assessment policy, the assignment of accommodations, and the use of accommodated scores in accountability policies that rate teachers, schools, and districts on their effectiveness. Related to validity is the desire for a test that is fair for all students, or that accurately measures what they know. *Fairness* is sometimes defined in the assessment literature as having equal levels of item difficulty between a referent group (in this case, students with access to spoken language) and students who are deaf or hard of hearing. The study of item difficulty discussed in the present review, Steinberg et al. (2009), indicated a mix of results as to the relative difficulty of accommodated versus unaccommodated test items for deaf and hard of hearing students. However, there was not enough specificity in either the sample demographics or the accommodations used to draw strong conclusions about the relative fairness of accommodated items for deaf and hard of hearing test takers. In the larger conversation about the potential changes in item constructs and the accessibility of test items, it would be necessary to look at item difficulty for an ASL- or signed language–accommodated test for students who are deaf or hard of hearing, preferably with a range of student-level

factors discussed in the literature review for the present article. Yet validity and fairness within accountability systems such as NCLB or teacher merit programs are only indirectly measured by item difficulty analyses. A comprehensive study of student-level factors, accommodations, resultant scores, and their use would require an extensive database across multiple states and a consistent policy of accommodations use. Although complex and far from complete, the Common Core Standards Assessments currently under development have the potential to render this question answerable (i.e., the question of the effects of accommodations on student scores for deaf and hard of hearing students and the items' accessibility for these students), or at least amenable to modeling using sophisticated statistical techniques (Beretvas, Cawthon, Lockhart, & Kaye, 2012).

Conclusion

Large-scale, standardized assessment is likely to remain a foundational part of how the U.S. education system, for both K–12 and postsecondary levels, measures student proficiency and progress toward academic goals. Many of the issues raised in the present article have parallels for other students from diverse backgrounds, including students with other disabilities, English Language Learners, students from migrant families, and students who have otherwise not had the opportunity to learn in high-quality instructional settings. Research into the specific implications of accommodations for students who are deaf or hard of hearing is largely inconclusive, at least if one considers the high standards of evidence required for a "gold standard." The studies discussed in the present review address many issues, yet, in many cases, only for a broadly defined student population and with-

out replication across settings and subject areas. For example, not much is known about the effects of accommodations for deaf and hard of hearing students who have additional disabilities, the range of quality in the implementation of ASL accommodations, or the resultant implications for test scores in how decisions are made about student proficiency and achievement. (For example, will the resultant score determine whether or not the student will receive a high school diploma?)

Yet even with these grey areas, accommodated assessments are being used every day to make high-stakes decisions that affect the future of students, teachers, and schools. In what is perhaps good news for inclusive assessment policies, in no instance in the present review did it appear that having an ASL-presented version of the assessment made the test items easier for students using the accommodation—a finding that minimizes concerns about test score validity. Future test systems, including those under development for the Common Core Standards, may afford further opportunities for increased access to test content for students who are deaf or hard of hearing. For example, although in the past using a computer to take a test may have been considered an accommodation, current assessment reform efforts will make it the norm. The move toward computer-based assessments may make it possible to individualize accommodations to address student-level and test-level characteristics in ways not available with traditional paper-and-pencil–based tests. We hope that this literature review motivates test developers and researchers alike to think critically about how assessment accommodations can facilitate learning outcomes for students who are deaf or hard of hearing.

References

Note. An asterisk (*) indicates an article included in the literature review.

Abedi, J. (2006). Language issues in item development. In S. Downing & T. Haladyna (Eds.), *Handbook of test development* (pp. 377–399). Rahway, NJ: Erlbaum.

Abedi, J. (2011). Language issues in the design of accessible items. In S. N. Elliott, R. J. Kettler, P. A. Beddow, & A. Kurz (Eds.), *Handbook of accessible achievement tests for all students: Bridging the gaps between research, practices, and policy* (pp. 217–230). New York, NY: Springer.

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

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *The standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Americans With Disabilities Act of 1990, Pub. L. No. 101-336, § 2, 104 Stat. 328 (1991).

*Anderson-Inman, L., Terrazas, F., & Slabin, U. (2008). *Supported Video Project: The use of expanded captions to promote student comprehension of educational videos on DVD.* Retrieved from National Center for Supported eText website: http://www.national-techcenter.org/documents/accessible_video_content_finalReport.pdf

*Ansell, E., & Pagliaro, C. M. (2006). The relative difficulty of signed arithmetic story problems for primary-level deaf and hard-of-hearing students. *Journal of Deaf Studies and Deaf Education, 11*(2), 153–170. doi:10.1093/deafed/enj030

Baker, E. (2010). *What probably works in an alternate assessment? What probably works in alternative assessment* (Report No. 772). Los Angeles: Center for Research on Evaluation Standards and Student Testing, University of California, Los Angeles.

Beretvas, S. N., Cawthon, S., Lockhart, L., & Kaye, A. (2012). Assessing impact, DIF and DFF in accommodated item scores: A comparison of multilevel measurement model parameterizations. *Educational and Psychological Measurement, 72*(5), 754–773. doi:10.1177/0013164412440998

Bielinski, J., Ysseldyke, J. E., Bolt, S., Friedebach, J., & Friedebach, M. (2001). Prevalence of accommodations for students with disabilities participating in a statewide testing program. *Assessment for Effective Intervention, 26*, 21–28.

Bolt, S. E., & Thurlow, M. L. (2004). Five of the most frequently allowed testing accommodations in state policy. *Remedial and Special Education, 25*(3), 141–152.

373

Bolt, S. E., & Thurlow, M. L. (2006). Item-level effects of the read-aloud accommodation for students with reading disabilities. *Assessment for Effective Intervention, 33,* 15–28.

Braden, J. P., & Hannah, J. M. (1998). Assessment of hearing-impaired and deaf children with the WISC-III. In A. Prifitera & D. Saklofske (Eds.), *WISC-III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 175–200). San Diego, CA: Elsevier Science.

Cawthon, S. (2001). Teaching strategies in inclusive classrooms with Deaf students. *Journal of Deaf Studies and Deaf Education, 6*(3), 212–225.

Cawthon, S. (2007). Hidden benefits and unintended consequences of No Child Left Behind polices for students who are Deaf or hard of hearing. *American Educational Research Journal, 44*(3), 460–492.

Cawthon, S. (2009). Making decisions about assessment practices for students who are Deaf or hard of hearing. *Remedial and Special Education, 32*(1), 4–21.

Cawthon, S. (2011). Test item linguistic complexity and assessments for deaf students. *American Annals of the Deaf, 156*(3), 255–269.

Cawthon, S., Beretvas, S. N., Lockhart, L., & Kaye, A. (2012). Factor structure of Opportunity to Learn for students with and without disabilities. *Educational Policy Analysis Archives, 20*(41). Retrieved from Arizona State University website: http://epaa.asu.edu/ojs/article/view/1043

Cawthon, S., Highley, K., & Leppo, R. (2011). Test item modifications for English Language Learners: Review of the empirical literature and recommendations for practice. *School Psychology Forum: Research in Practice, 5*(2), 28–41.

Cawthon, S., Ho, E., Patel, P., Potvin, D., & Trundt, K. (2009). Toward a multiple construct model of measuring the validity of assessment accommodations. *Practical Assessment, Evaluation, and Research, 14*(21). Retrieved from http://pareonline.net/genpare.asp?wh=0&abt=14

Cawthon, S., Kaye, A., Lockhart, L., & Beretvas, S. N. (2012). Effects of linguistic complexity and accommodations on estimates of ability for students with learning disabilities. *Journal of School Psychology, 50,* 293–316.

Cawthon, S., Leppo, R., Carr, T., & Kopriva, R. (2013). Toward accessible assessments: The promises and limitations of test item adaptations for students with disabilities and English Language Learners. *Educational Assessment, 18*(2), 73–98. doi: 10.1080/10627197.2013.789294

Cawthon, S., & Online Research Lab. (2006). Findings from the National Survey on Accommodations and Alternate Assessments for Students Who Are Deaf or Hard of Hearing. *Journal of Deaf Studies and Deaf Education, 11*(3), 337–359.

Cawthon, S., & Online Research Lab. (2007). Accommodations use for statewide standardized assessments: Prevalence and recommendations for students who are deaf or hard of hearing. *Journal of Deaf Studies and Deaf Education, 13*(1), 55–96.

*Cawthon, S., Winton, S., Garberoglio, C., & Gobble, M. (2011). The effects of American Sign Language as an assessment accommodation for students who are deaf or hard of hearing. *Journal of Deaf Studies and Deaf Education, 16*(2), 198–211.

Cawthon, S., Wurtz, K., & Online Research Lab. (2010). Predictors of assessment accommodations use for students who are deaf or hard of hearing. *Journal of Educational Research and Policy Studies, 10*(1), 17–35.

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 No. 83). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Christensen, L. L., Lazarus, S. S., Crone, M., & Thurlow, M. L. (2008). *2007 state policies on assessment participation and accommodations for students with disabilities* (Synthesis Report No. 69). Minneapolis: 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 No. 56). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Common Core State Standards Initiative. (2010). Common core state standards for English language arts and literacy in history/social studies, science, and technical subjects, Retrieved from http://corestandards.org/assets/CCSSI_ELA%20Standards.pdf

Crawford, L., & Tindal, G. (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*(2), 71–88.

Czubek, T. A. (2006). Blue Listerine, parochialism, and ASL literacy. *Journal of Deaf Studies and Deaf Education, 11*(3), 373–381. doi:10.1093/deafed/enj033

Easterbrooks, S. R., Lederberg, A. R., & Connor, C. M. (2010). Contributions of the emergent literacy environment to literacy outcomes for young children who are deaf. *American Annals of the Deaf, 155*(4), 467–480. doi: 10.1353/aad.2010.0024

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

Elliott, S. N., & Roach, A. T. (2007). Alternate assessments in students with significant disabilities: Alternative approaches, common

technical challenges. *Applied Measurement in Education, 20,* 301–333.

Emmorey, K., Petrich, J. A., & Gollan, T. H. (2013). Bimodal bilingualism and the frequency-lag hypothesis. *Journal of Deaf Studies and Deaf Education, 18*(1), 1–11. doi:10.1093/deafed/ens034

Fairbairn, S., & Fox, J. (2009). Inclusive achievement testing for linguistically and culturally diverse test takers: Essential considerations for test developers and test makers. *Educational Measurement: Issues and Practice, 28*(1), 10–24.

Fletcher, J., 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,* 136–150.

*Hoffman, M., & Wang, Y. (2010). The use of graphic representations of sign language in leveled texts to support deaf readers. *American Annals of the Deaf, 155*(2), 131–136.

Individuals With Disabilities Education Improvement Act of 2004, Pub. L. No. 108-446, 118 Stat. 2647.

*Johnson, E., Kimball, K., & Brown, S. (2001). American Sign Language as an accommodation during standards-based assessments. *Assessment for Effective Intervention, 26,* 39–47. doi: 10.1177/073724770102600207

Keiffer, M. J., Lesaux, N. R., Rivera, M., & Francis, D. J. (2009). Accommodations for English Language Learners taking large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research, 79*(3), 1168–1201.

Kent, B., & Smith, S. (2006). They only see it when the sun shines in my ears: Exploring perceptions of adolescent hearing aid users. *Journal of Deaf Studies and Deaf Education, 11*(4), 461–476. doi: 10.1093/deafed/enj044

Kopriva, R. J. (2008). *Improving testing for English Language Learners: A comprehensive approach to designing, building, implementing, and interpreting better academic assessments.* New York, NY: Routledge.

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

Kyle, F. E., & Harris, M. (2011). Longitudinal patterns of emerging literacy in beginning deaf and hearing readers. *Journal of Deaf Studies and Deaf Education, 16*(3), 289–304. doi:10.1093/deafed/enq069

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 No. 64). Minneapolis: University of Minnesota, National Center on

Educational Outcomes. Retrieved from University of Minnesota website: http://education.umn.edu/NCEO/OnlinePubs/Synthesis64/

Liu, K. K., & Anderson, M. (2008). Universal Design considerations for improving student achievement on English language proficiency tests. *Assessment for Effective Intervention, 33,* 167–176.

Luckner, J. L., & Bowen, S. (2006). Assessment practices of professionals serving students who are deaf or hard of hearing: An initial investigation. *American Annals of the Deaf, 151*(4), 410–417.

Luetke-Stahlman, B., & Nielsen, D. C. (2003). The contribution of phonological awareness and receptive and expressive English to the reading ability of deaf students with varying degrees of exposure to accurate English. *Journal of Deaf Studies and Deaf Education*, *8*(4), 464–484.

*Maihoff, N., Bosso, E., Zhang, L., Fischgrund, J., Schulz, J., Carlson, J., & Carlson, J. E. (2000). *The effects of administering an ASL signed standardized test via DVD player/television and by paper-and-pencil: A pilot study*. Dover: Delaware Department of Education.

Mandinach, E., Bridgeman, B., Cahalan-Laitusis, C., & Tripani, C. (2005). *The impact of extended time on SAT test performance.* New York, NY: College Board.

*Marschark, M., Convertino, C., Sapere, P., & Seewagen, R. (2006). Access to postsecondary education through sign language interpreting. *Journal of Deaf Studies and Deaf Education, 10*(1), 38–50.

*Marschark, M., Pelz, J., Convertino, C., Sapere, P., Arndt, M. E., & Seewagen, R. (2005). Classroom interpreting and visual information processing in mainstream education for deaf students: Live or Memorex®? *American Education Research Journal*, *42*(2), 727–761.

*Marschark, M., Sapere, P., Convertino, C., Mayer, C. Wauters, L., & Sarchet, T. (2009). Are deaf students' reading challenges really about reading? *American Annals of the Deaf, 154*(4), 357–370.

*Marschark, M., Sapere, P., Convertino, C., & Pelz, J. (2008). Learning via direct and mediated instruction by deaf students. *Journal of Deaf Studies and Deaf Education 13*(4), 546–561. doi: 10.1093/deafed/enn014

Mayer, C. (2007). What really matters in the early literacy development of deaf children. *Journal of Deaf Studies and Deaf Education*, *12*(4), 411–431. doi:10.1093/deafed/enm020

Mayer, C., & Akamatsu, C. (1999). Bilingual-bicultural models of literacy education for deaf students: Considering the claims. *Journal of Deaf Studies and Deaf Education*, *4*(1), 1–8. doi: 10.1093/deafed/4.1.1

McIlroy, G., & Storbeck, C. (2011). Development of deaf identity: An ethnographic study. *Journal of Deaf Studies and Deaf Education*, *16*(4), 494–511. doi:10.1093/deafed/enr017

Meloy, L. L., Deville, C. & Frisbie, D. A. (2002). The effects of a read-aloud accommodation on test scores of students with and without a learning disability in reading. *Remedial and Special Education, 23*(4)*,* 248–255.

Mitchell, R. (2004). National profile of deaf and hard of hearing students in special education from weighted survey results. *American Annals of the Deaf, 149*(4), 336–349. doi: 10.1093/deafed/enj004

Mitchell, R. (2006). How many deaf people are there in the United States? Estimates from the survey of income and program participation. *Journal of Deaf Studies and Deaf Education*, *11*(1), 112–119.

Mitchell, R., Qi, S., & Traxler, C. B. (2007). *Stanford Achievement Test, 10th edition, national performance norms for deaf and hard of hearing students: A technical report.* Washington, DC: Gallaudet University, Gallaudet Research Institute.

Moores, D. F. (2010). Integration inclusion oblivion. *American Annals of the Deaf*, *155*(4), 395–396. doi:10.1353/aad.2010.0026

*Mowl, H. (1985). *Performance of deaf students on a standard and modified minimum competency test* (Unpublished doctoral dissertation). University of Pittsburgh, Pittsburgh, PA. Retrieved from University of Texas Libraries website: http://ezproxy.lib.utexas.edu/login?url=http://search.proquest.com.ezproxy.lib.utexas.edu/docview/303398077?accountid=7118

Myers, C., Clark, M. D., Musyoka, M. M., Anderson, M. L., Gilbert, G. L., Agyen, S., & Hauser, P. C. (2010). Black deaf individuals' reading skills: Influence of ASL, culture, family characteristics, reading experience, and education. *American Annals of the Deaf*, *155*(4), 449–457. doi:10.1353/aad.2010.0044

Nikolaraizi, M., & Hadjikakou, K. (2006). The role of educational experiences in the development of deaf identity. *Journal of Deaf Studies and Deaf Education*, *11*(4), 477–492. doi: 10.1093/deafed/enl003

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 U.S.C. § 1425 (2002).

Padden, C., & Ramsey, C. (1993). Deaf culture and literacy. *American Annals of the Deaf, 138*(2), 96–99. doi: 10.1353/aad.2012.0623

Paul, P. V. (2006). New literacies, multiple literacies, unlimited literacies: What now, what next, where to? A response to "Blue Listerine, parochialism, and ASL literacy." *Journal of Deaf Studies and Deaf Education*, *11*(3), 382–387. doi: 10.1093/deafed/enj039

Qi, S., & Mitchell, R. E. (2012). Large-scale academic achievement testing of deaf and hard-of-hearing students: Past, present, and future. *Journal of Deaf Studies and Deaf Education 17*(1), 1–18. doi:10.1093/deafed/enr028

Rigney, S., Wiley D. E., and Kopriva, R. J. (2008). The past as preparation: Measurement, public policy, and implications for access. In R. J. Kopriva (Ed.), *Improving testing for English Language Learners: A comprehensive approach to designing, building, implementing, and interpreting better academic assessments* (pp. 39–66). New York, NY: Routledge.

*Russell, M., Kavanaugh, M., Masters, J., Higgins, J., & Hoffmann, T. (2009). Computer-based signing accommodations: Comparing a recorded human with an avatar. *Journal of Applied Testing Technology*, *10*(3). Retrieved from Association of Test Publishers website: http://www.testpublishers.org/assets/documents/computer_based.pdf

Schick, B., Skalicky, A., Edwards, T., Kushalnagar, P., Topolski, T., & Patrick, D. (2013). School placement and perceived quality of life in youth who are deaf or hard of hearing. *Journal of Deaf Studies and Deaf Education*, *18*(1), 47–61. doi:10.1093/deafed/ens039

Schulte, A. G., Elliot, S. N., & Kratochwill, T. R. (2001). Effects of testing accommodations on standardized mathematics test scores: An experimental analysis of the performances of students with and without disabilities. *School Psychology Review, 30*(4), 527–547.

Shaftel, J., Belton-Kocher, E., Glassnap, D., & Poggio, J. (2006) The impact of language characteristics in mathematics test items on the performance of English Language Learners and students with disabilities. *Educational Assessment*, *11*(2), 105–126. doi: 10.1207/s15326977ea1102_2

Shaftel, J., Yang, X., Glassnap, D., & Poggio, J. (2005). Improving assessment validity for students with disabilities in large-scale assessment programs. *Educational Assessment, 10*(4), 357–375. doi: 10.1207/s15326977ea1004_3

Singleton, J. L., Morgan, D., DiGello, E., Wiles, J., & Rivers, R. (2004). Vocabulary use by low, moderate, and high ASL-proficient writers compared to hearing ESL and monolingual speakers. *Journal of Deaf Studies and Deaf Education*, *9*(1), 86–103. doi:10.1093/deafed/enh011

Sireci, S. G., Li, S., & Scarpati, S. (2003). *The effects of test accommodations on test performance: A review of the literature* (Center for Educational Assessment Research Report No. 485). Amherst: University of Massachusetts, School of Education.

Sireci, S. G., Scarpati, S., & Li, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. *Review of Educational Research, 75*(4), 457–490. doi:10.3102/00346543075004457

Smiler, K., & McKee, R. L. (2007). Perceptions of Maori deaf identity in New Zealand. *Journal of Deaf Studies and Deaf Education*, *12*(1), 93–111. doi:10.1093/deafed/enl023

Soukup, M., & Feinstein, S. (2007). Identification, assessment, and intervention strategies for deaf and hard of hearing students with learning disabilities. *American Annals of the Deaf, 152*(1), 56–62.

Spencer, L. J., Tomblin, J. B., & Gantz, B. J. (2012). Growing up with a cochlear implant: Education, vocation, and affiliation. *Journal of Deaf Studies and Deaf Education, 17*(4), 483–498. doi:10.1093/deafed/ens024

Spencer, P. E., & Marschark, M. (2010). *Evidence-based practice in educating deaf and hard of hearing students*. New York, NY: Oxford University Press.

*Steinberg, J., Cline, F., Ling, G., Cook, L., & Tongatta, N. (2009). *Examining the validity and fairness of a state standards-based assessment of English-language arts for deaf or hard of hearing students*. Princeton, NJ: Educational Testing Service.

*Stinson, M., Elliot, L., Kelly, R., & Liu, Y. (2009). Deaf and hard-of-hearing students' memory of lectures with speech-to-text and interpreting/note-taking services. *Journal of Special Education, 43*, 52–64.

Strong, M., & Prinz, P. M. (1997). A study of the relationship between American Sign Language and English literacy. *Journal of Deaf Studies and Deaf Education*, *2*(1), 37–46.

*Sullivan, P., & Schulte, L. (1992). Factor analysis of WISC-R with deaf and hard-of-hearing children. *Psychological Assessment, 4*(4), 537–540.

Thurlow, M. L., Lazarus, S., Thompson, S., & Robey, J. (2002). 2001 state policies on assessment participation and accommodations (Synthesis Report No. 46). Minneapolis: University of Minnesota, National Center on Educational Outcomes. Retrieved from University of Minnesota website: http://education.umn.edu/NCEO/OnlinePubs/Synthesis46.html

Towles-Reeves, E., Kleinert, H., & Muhomba, M. (2009). Alternate assessment: Have we learned anything new? Exceptional Children, 75, 233–252.

Vetter, A., Löhle, E., Bengel, J., & Burger, T. (2010). The integration experience of hearing impaired elementary school students in separated and integrated school settings. *American Annals of the Deaf*, *155*(3), 369–376. doi:10.1353/aad.2010.0015

Wilbur, R. B. (2000). The use of ASL to support the development of English and literacy. *Journal of Deaf Studies and Deaf Education*, *5*(1), 81–104. doi: 10.1093/deafed/5.1.81

Williams, C. (2012). Promoting vocabulary learning in young children who are d/Deaf and hard of hearing: Translating research into practice. *American Annals of the Deaf 156*(5), 501–508. doi:10.1353/aad .2012.1597

*Wolf, J. (2007). *The effects of testing accommodations usage on students' standardized test scores for deaf and hard-of-hearing students in Arizona public schools* (Unpublished doctoral dissertation). University of Arizona, Tucson.

Zatta, M., & Pullin, D. (2004). Education and alternate assessment for students with significant cognitive disabilities: Implications for educators. *Education Policy Analysis Archives, 12*(16). Retrieved from Arizona State University website: http://epaa.asu .edu/epaa/v12n16/

Zenisky, A. L., & Sireci, S. G. (2007). *A summary of the research on the effects of test accommodations: 2005–2006* (Technical Report No. 47). Minneapolis: University of Minnesota, National Center on Educational Outcomes. Retrieved from University of Minnesota website: http://cehd.umn.edu/nceo/OnlinePubs/ Tech47/default.html