Science and Evidence of Success: Two Emerging Issues in Assessment Accommodations for Students Who Are Deaf or Hard of Hearing

Stephanie W. Cawthon
The University of Texas at Austin

In the United States, students who are deaf or hard of hearing (SDHH) are required to participate in high-stakes standardized assessments under No Child Left Behind reforms. In 2006–2007, states added science to reading and mathematics as a tested content area. Many SDHH participate in these assessments using testing accommodations, but teachers have few evidence-based resources to draw upon when making accommodations decisions. Two research questions guided this study: (a) What were patterns of SDHH 2006–2007 test accommodations use in state standardized assessments in mathematics, reading, and science? (b) What evidence did teachers use to determine the effectiveness of accommodations for SDHH? A total of 290 participants described their assessment practices with SDHH via an online and paper-and-pencil survey. Extended time, small group, and test directions interpreted were the most frequently used accommodations by SDHH, but there were some different patterns in science for accommodations that included changes to the test items. Teachers reported both student satisfaction and test score validity epistemologies of accommodations’ effectiveness.

The role of assessment in educational reforms has heightened greatly over the past two decades. In the current accountability reform movement, including the No Child Left Behind Act (NCLB) of 2001, assessments serve as the primary indicator of academic success and evidence of school progress toward closing the achievement gap. Standardized assessments, those that use a paper-and-pencil format and administer the same set of items to all participants, are the most common approach to measuring student proficiency under accountability reforms. Students, however, come to the assessment context with diversity in their language, disability, and academic preparation (Abedi, 2003). Some of these characteristics are not a good match for a standardized assessment format (Phillips, 1994). Test accommodations are often used to help students gain access to the standardized test format in order to allow students to show their true content area knowledge and skill (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999; Thurlow & Bolt, 2001). In this way, students with disabilities or who are English language learners are able to participate more meaningfully in the accountability reform process.

Students who are deaf or hard of hearing (SDHH) are often among those students with diverse language, disability, and academic preparation characteristics (Marschark, Lang, & Albertini, 2002). Test accommodations play a key role in how SDHH are included in the high-stakes standardized assessment process. Previous research indicates that SDHH use a wide range of accommodations, including extended time (ET), separate setting for test administration, test directions interpreted (TDI), test items interpreted (TII), and test items read aloud (TIR; Cawthon & the Online Research Lab, 2006, 2007). The purpose of this article is to explore two emerging issues in the practice of assessment accommodations for SDHH: (a) science as a new content area and (b) sources of evidence for the effectiveness of accommodations. A brief summary of the rationale for each research focus is provided below.
NCLB recently broadened the range of content areas that were to be included on state assessments. In 2006–2007, states were required to add science to reading and mathematics as a tested content area in accountability reform. In the previous research, educators reported that accommodations that changed the presentation of the test items, such as reading the test question aloud or having it signed to the student, were used less frequently for reading than for mathematics (Cawthon & the Online Research Lab, 2007). Science is not only a new content area, one that is highly dependent on vocabulary knowledge, similar to reading, but also involves problem-solving and quantitative skills, as used for mathematics (WestEd & Council of Chief State School Officers [CCSSO], 2005). To date, we do not know whether the patterns of accommodations use for SDHH in science are more similar to those used for reading or for mathematics, or neither subject area.

As with students with disabilities as a whole, many decisions about assessment accommodations for SDHH are made in the absence of comprehensive data on their effectiveness. Individualized Education Program (IEP) team members rarely have information about whether or not the recommended accommodations increased access to test content (Elliott, McKevitt, & Kettler, 2002; Schulte, Elliott, & Kratochwill, 2001). Teachers do, however, have access to potential sources of evidence such as feedback from the student, comparison with other students using the accommodations, or the resulting test score (e.g., Cawthon, 2009; Fuchs & Fuchs, 2001; Kopriva, Emick, & Hipolito-Delgado, 2007). This study was designed to explore epistemologies of teachers making accommodations decisions in two ways. First, to identify what sources of information teachers used when evaluating the effectiveness of a test accommodation for SDHH and, second, to gather information about what issues they thought affected the validity of an accommodated test score for SDHH.

To provide sufficient context for these research foci, this article first presents background information on SDHH and important characteristics that influence how they participate in standardized assessment. The article then describes the current literature on assessment accommodations for SDHH, what is known about accommodations for mathematics and reading, and potential considerations when using accommodations for science tests. The literature review concludes with information about teacher epistemology and accommodations decision making, for students with disabilities as a whole and for SDHH, specifically.

Students Who Are Deaf or Hard of Hearing

SDHH form a heterogeneous group with a wide variety of experiences with language both at home and at school (Mitchell, 2004; Van Cleve & Crouch, 1989). The starting point for some families of children who are deaf or hard of hearing is the degree of the child’s hearing loss (Marschark & Spencer, 2006). Hearing loss is defined as ranging from mild to profound, is defined in terms of “decibel loss,” and can occur in one or both ears (Vohr, 2003). Some types of hearing loss are inherited and carry from generation to generation, whereas other types are due to conditions such as otitis media or a traumatic head injury (Easterbrooks, 1999). Timing of onset can also vary: some losses have an onset at birth and some losses change as the child grows older. In recent years, early hearing screening of babies at birth has resulted in a larger proportion of children with hearing loss identified at a very young age (National Center for Hearing Assessment and Management, 2008; Yoshinaga-Itano & Sedey, 2000). Once a hearing loss is identified, parents have a range of options that vary in their emphasis on American Sign Language (ASL) and/or on speech and amplified hearing.

SDHH are a unique group in that there are both sensory and cultural factors that affect their linguistic and, by extension, educational development (Loeterman, Paul, & Donahue, 2002; Marschark et al., 2002). For all children, a student’s academic success depends, in part, on the language and literacy development support the child receives in early and middle childhood (Antia, Jones, Reed, & Kreimeyer, 2009; Johnson, Liddell, & Erting, 1989). Choices about language use are not simply about the degree of hearing loss but also about the home language of the family members (Lucas & Valli, 1992; Mitchell & Karchmer, 2004). For example, in families with deaf parents who use ASL as the home language, all children in the family will be exposed to ASL, even
children who do not have a hearing loss (Mitchell & Karchmer, 2004). Yet without an outside intervention or parent ASL language program, children with a severe or profound hearing loss who have hearing parents may not have exposure to ASL in their very early months and years (Padden & Humphries, 2005). The interaction between individual factors and family factors results in a variety of language and, as a result, a range of academic development paths for SDHH (Antia, Sabers, & Stinson, 2007; Powers, 2003).

Assessment Accommodations for SDHH

All students, including SDHH, are to be included in state assessments for high-stakes decision making (NCLB, 2002). High-stakes nature of assessment refers to the use of test scores for decisions such as a student’s progress or a school’s effectiveness. For SDHH, particularly those with diverse home and school language experiences, participating in standardized assessments frequently requires the use of test accommodations (Cawthon & the Online Research Lab, 2006, 2007; Cawthon, Hersh, Kim, and the Online Research Lab, 2009). The next section of the article summarizes the kinds of accommodations SDHH use and some differences in how they are used for mathematics and reading content areas. This section will conclude with thoughts about science and the potential issues that may arise when implementing accommodations for this new content area.

Prevalence

Choices about accommodations are regulated by state assessment policies for students with disabilities as well as guidelines for assessment under the Individuals with Disabilities Education Improvement Act (IDEIA, 2004). On the one hand, assessment decision making is to focus on students’ individual needs. On the other hand, state policies are meant to ensure that test scores are valid and reliable across both accommodated and unaccommodated tests. There are six accommodations for which prevalence has been measured over the past 30 years: ET, small group/individual administration (SGI), TDI, TIR, TII, and student signs response (SSR; Cawthon & the Online Research Lab, 2006, 2007). A description of these accommodations is provided in Table 1. For SDHH, the most prevalent accommodations are ET, SGI, and TDI. Less frequently used but still important options noted by many teachers and assessment personnel who work with SDHH are TII and TIR presentation of test items. Only a small proportion of SDHH uses SSR; this low prevalence rate reflects the fact that state policies largely restrict this accommodations option (Bolt & Thurlow, 2004). The restrictions are based on a range of factors, including staffing needs (i.e., a trained scribe to translate student signs into written form) and questions about the accuracy of translations from student signs to the scribe’s written documentation of a student’s responses.

Mathematics Versus Reading

The validity of an accommodated test score depends, in part, on the content area of the assessment (Sireci, Scarpati, & Li, 2005). State policies on accommodations use are designed to allow students the greatest access to the test content while minimizing threats to the interpretability of the test score (Bolt & Thurlow, 2004). In other words, can the test score be seen as “equal” to an unaccommodated test score in terms of the knowledge or skill it represents? If a test accommodation makes the test item more or less difficult than an unaccommodated item, the resulting test scores are no longer comparable. Efforts to increase access to test content through accommodations can sometimes pose potential risks to the validity of the test scores (Elliott et al., 2002). This is particularly true when the accommodation removes part of the task that the test is trying to measure (Crawford, Helwig, & Tindal, 2004; Fletcher et al., 2006). Take, for example, a mathematics test where the goal is to measure how well a student knows her multiplication tables. If the student uses a calculator as an accommodation on that assessment, the test score no longer represents her ability to multiply numbers (but rather, her ability to use a calculator). It is important, therefore, to be aware of the accommodation match with test area content.

The match between accommodation and test area content holds special significance for SDHH. As has been true for many years, SDHH often face delays in
their literacy development (see Antia et al., 2009, for a recent review of academic achievement trends). There are several potential reasons for this, including early delays in exposure to language or, for emerging readers, difficulty in decoding written English without access to the phonetic repertoire of spoken English. Standardized, norm-referenced tests of reading ability highlight the struggles that deaf education models have had in bringing SDHH to grade level in reading (e.g., Traxler, 2000). Delays in reading pose significant challenges not only in language arts but also in developing knowledge in language-rich content areas. Some assessment accommodations attempt to remove reading skills as a barrier to accessing test content. For mathematics tests, an accommodation for SDHH may include having test items read aloud or translated into ASL (or other sign system), particularly for word problems that rely on language-embedded content. There is potentially less concern about validity for an accommodated word problem than there might be for the calculator accommodation–multiplication table task described above. Yet even for mathematics problems, translated test items may change the problem type (Ansell & Pagliaro, 2001). The difficulty arises

<table>
<thead>
<tr>
<th>Accommodation type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extended time</td>
<td>On a timed test, the student may receive more time to complete the assessment. Depending on the context and the student’s needs, “extended” may mean time and a half, double time, or as long as the student needs to complete the test items. This accommodation does not apply to assessments that are designed to be untimed for all students.</td>
</tr>
<tr>
<td>Small group/individual administration (SGI)</td>
<td>In a standard test setting, students often take the assessment as a classroom or large group. An SGI of an assessment is an accommodation designed to either (a) reduce distraction or (b) make it possible to administer other accommodations without distracting other test takers. This accommodation is often implemented in a separate room or individual carrel.</td>
</tr>
<tr>
<td>Test directions interpreted (TDI)</td>
<td>In a standard test setting, test directions are often read out loud by the test administrator. For students whose first language is not spoken English, test directions in this manner may place them at a disadvantage for success on the assessment. For SDHH, TDI may mean that they are provided in ASL or in another visual communication mode. This accommodation is often implemented by an individual familiar to the student to facilitate communication and understanding of test directions.</td>
</tr>
<tr>
<td>Test items read aloud (TIR)</td>
<td>In a standard test setting, students read test items to themselves, silently, and then respond with their answers written on a bubble sheet (for multiple-choice questions). In some cases, students have difficulty decoding or comprehending the test item. The TIR accommodation is when a test administrator reads the test to the student, who then responds with their answer on the bubble sheet.</td>
</tr>
<tr>
<td>Test items interpreted (TII)</td>
<td>The TII is similar to the TIR accommodation, above. The TII accommodation involves the test administrator translating the text of the test item into ASL or other sign system for the SDHH. At times, a student may receive a combination of TIR and TII, depending on their language preferences and the test context.</td>
</tr>
<tr>
<td>Student signs response (SSR)</td>
<td>SSR focuses on the student’s method of providing answers to the test items. There may be reasons that recording an answer on a bubble sheet is difficult for a student or, for student-constructed items, that it is difficult for a student to express himself/herself in writing. In these cases, a student might use ASL or other sign system to provide their responses. The test administrator or a scribe would then translate the response into written English or record the chosen bubble response.</td>
</tr>
</tbody>
</table>
when the test item is meant to measure reading abilities and the accommodation (i.e., TII and TIR) removes reading as a barrier to the test content. Because of the complexity of reading as a skill, there can be quite a bit of murkiness in the accommodation task match for reading. Reading involves many subcomponents ranging from vocabulary knowledge to fluency to reading comprehension and interpretation (Schirmer, 2000). At what point does physically reading (i.e., decoding) the words on the page become a necessary component of answering the test item? In turn, at what point does having an accommodation that removes this process from answering the test item render the resulting test score invalid? From a practical standpoint, it may not be possible to dissect a test into items that can and cannot be used with TII or TIR as an accommodation. In this case, assessment teams may decide to over accommodate (use TII or TIR for the entire reading test) or under accommodate (not allow the accommodation at all).

Science as a New Content Area

The addition of science as a new content area in the required NCLB assessments poses not only a new opportunity for SDHH to participate in accountability reform but also some challenges (NCLB, 2002). Science as a field has some similarities and differences with mathematics and reading. A similarity with mathematics is that there is content area–specific vocabulary (i.e., hypotenuse and oxidation) that can be intrinsic to understanding the test item demands (WestEd & CCSSO, 2005). This is particularly true when comparing mathematics word problems with science problems (as opposed to computation problems such as the multiplication example above). With field-specific vocabulary such as what is found in science, even if one can read (decode) the word, the student must know what the term means to be able to answer questions about those concepts (Stansfield, 2002). In contrast with an accommodation such as ET (Kim, Schneider, & Siskind, 2009), providing a student with the definition of the word may also provide extra assistance in completing the item itself. In ASL, vocabulary that is fingerspelled (e.g., h-y-p-o-t-e-n-u-s-e) will provide less information than a word that is translated using its sign. A central question here is whether a test item using the spelled out versions of vocabulary is easier or harder than the one using the ASL sign or visual translation. As a content area full of procedural and conceptual vocabulary, to what extent does translation of test items change the nature of what is being measured? And, by extension, the validity of the accommodated test score?

One difference between science and other fields is its complexity, even in very early grades. Depending on the objectives and content area standards measured, a single science assessment, or even test item, may tap into a student’s content knowledge, problem-solving ability, vocabulary, laboratory experience, and ability to apply information (Patz, 2005). The format of test items in science may be an important factor in how students participate and thus how students use test accommodations. For example, in the 2009 National Assessment of Educational Progress, only half of the items were in a multiple-choice format (WestEd & CCSSO, 2005). The remaining items were student-constructed response items, hands-on problem solving, and computer-interactive tasks. These item formats are meant to allow students to display the kind of higher order thinking and experimental procedural knowledge inherent in science as a field.

To illustrate this point about item formats and test item demands, consider the following two examples of science test items one might find on a middle school assessment.

Item Example 1

The surface of the Moon is covered with craters. Most of these craters were formed by …

1. eruptions of active volcanoes,
2. the impacts of many meteoroids,
3. shifting rock on the Moon’s surface (“moonquakes”), and
4. tidal forces caused by the Earth and the Sun.

Item Example 2

List four ways that the Earth is different from the Moon.
The first example parallels test items on many standardized assessments and focuses on concept knowledge. The second example also focuses on concept knowledge but uses a different approach, a student-constructed response, to understand what the student knows about the relationship between the Earth and the Moon. Translating the first test item into ASL would involve visual representations of the word “crater.” If fingerspelled, as in c-r-a-t-e-r, there would be little additional information provided to the student. If signed as the ASL sign for crater, the student would be given some additional clues about the physical properties of craters that might lead him to pick one of the responses that could be inferred by the sign. For the second item, in contrast, an ASL-translated test item would not provide additional information or provide links between the question stem and the responses (because the student constructs the responses herself). Open-ended questions thus allow for lessened threats to the validity of the test score because there are fewer connections made for the student between the prompt and student response. Because of the nature of science tasks and the emphasis on student-constructed test items, there may be a greater potential for fair use of TIR and TII accommodations in this subject area.

The level of reading skill needed to complete science items, and the extent to which reading accommodations interfere with test item content, varies tremendously from item to item (Council of Chief State School Officers, 2008). In addition to field-specific vocabulary, science tests can also include demonstrations of procedural knowledge (Patz, 2005). Procedural knowledge tasks tend to have lower reading demand than factual knowledge and recall. Two of the most common ways to capture procedural knowledge are concept maps and laboratory demonstrations. In a concept map, students quite literally draw a visual map of how concepts relate to each other, drawing nodes and connecting arrows with notation to describe the relationships between items (Test Item Example 3, below). In a laboratory demonstration, the student may be presented with a set of stimuli (e.g., battery, wire, and a light bulb) and asked to show how to light the light bulb and to describe what scientific principles explain the process (e.g., conservation of energy).

Test Item Example 3
Fill in the relationships between the following elements

In both examples of procedural test items, the test scoring process interprets what the student conveys through their model and explanation of a scientific relationship or property. In a sense, this kind of test item is similar to a student-constructed response (test example 2) but involves additional visual components and concrete examples to manipulate and use to demonstrate content knowledge (Ruiz-Primo, Schultz, Li, & Shavelson, 2001). This difference in response mode may result in different opportunities for accommodations as well. For example, once the test item is explained to the student, the most appropriate accommodation for a procedural task may be SSR or having a scribe document what the student conveys in their test response. Where possible, a digital recording of the student’s response may allow for a full understanding of what the student conveys during the test process. This approach is quite different than the policies and practices for students completing multiple-choice reading and mathematics items. Science as a content area may therefore provide new challenges but also opportunities for SDHH to participate in assessments using accommodations that do not threaten the validity of the resulting test score (Williams, Ebert-May, Luckie, Hodder, & Koptur, 2004).

Accommodations Decision Making
Decisions about accommodations, particularly for high-stakes testing, are challenging and complex. A
major component of the research literature around accommodations for standardized assessment centers on the decision-making process (Elliott et al., 2002; Schulte et al., 2001). In addition to the focus on meeting an individual student’s needs, decisions about accommodations are guided by district, state, and federal policy. Research on decision making has illustrated how difficult it is to meet multiple goals of test validity and student’s access to test content. In general, teachers tend to over accommodate or assign accommodations that either are not effective or are not allowed under state policies (Helwig & Tindal, 2003; Ketterlin-Geller, Alonzo, Braun-Monegan, & Tindal, 2007; Shriner & DeStefano, 2003). Decisions about accommodations have also been found to be inconsistent, meaning that students with the same characteristics may receive different accommodations, even if assigned by the same individual or IEP team. These findings have raised concerns about the decision-making process and encouraged the field to more clearly articulate criteria for test accommodations.

Decision Making for SDHH

Preliminary research on the decision-making process for SDHH revealed some important information about the priorities teachers have when assigning accommodations to their students (Cawthon, 2009; Cawthon & the Online Research Lab, 2007). First, teachers honor the intent of IDEIA not to categorically assign the same accommodations to all SDHH but to take each student’s individual circumstances into account. In other words, teachers largely assign accommodations on a case-by-case basis. Second, teachers are sensitive to the subject of the test when recommending accommodations. Teachers were more likely to recommend accommodations when a student was strong in the subject area (i.e., mathematics or reading) and more likely to recommend an alternate assessment (i.e., not the standardized assessment) when the student was weak in the subject area. Reading proficiency did not, contrary to expectations, predict accommodations choices for assessments outside of reading. In other words, student proficiency in reading only predicted accommodations choices for reading as its own subject area, not as a means to access content in other subject areas. Given the complexity of science content and assessments, decisions about accommodations for the new science subject area are important to explore.

Need for Evidence-Based Practice

An emphasis on evidence-based practice is one of the hallmarks of NCLB accountability reforms (NCLB, 2002). Evidence-based practice is a term that refers to a teaching or assessment approach that has shown to be effective for use by research that uses rigorous empirical design. The assumption behind this criterion is that evidence-based educational practice will lead to better student outcomes than practices that have not been substantiated in the research literature (e.g., interventions that are vetted by the “What Works Clearinghouse” at http://ies.ed.gov/ncee/wwc/) Assessment accommodations decision making largely lacks evidence-based models and procedures (Fuchs & Fuchs, 2001). The great variety of assessment policies from state to state reflects the unclear findings about evidence-based practice for testing students with disabilities (Clapper, Morse, Lazarus, Thompson, & Thurlow, 2005). Many teachers and IEP teams do not have set rubrics or decision paths that incorporate the multiple factors they must consider at their disposal (Edgemon, Jablonski, & Lloyd, 2006). For SDHH, teachers must largely make decisions in the absence of research of “what works” for assessment. Without a research base as a guide, teachers must consider local sources of information about the effects of accommodations they recommend for their SDHH.

Research Questions

The literature review provides an overview of three overlapping content areas that support the premise of this study: SDHH demographics, content area of assessment (including science), and assessment accommodations decision making converge to form the context of this current line of research. This study is an extension of accommodations research into a new subject area (science) as well as a further probe into the nature of evidence-based practice used by professionals in the field. Two research questions guided this
study: (a) What were patterns of SDHH 2006–2007 test accommodations use in state standardized assessments in mathematics, reading, and science? (b) What evidence did teachers use to determine the effectiveness of accommodations for SDHH?

Methods

The data for this article are drawn from the Fourth Annual National Survey of Assessments and Accommodations for Students Who Are Deaf or Hard of Hearing (Fourth National Survey). The Fourth National Survey ran from November 2007 through June 2008. Participants were asked to describe their assessment practices with SDHH during the 2006–2007 school year (the year prior to the survey).

Sample

The sample for this study consisted of teachers or other educational professionals who worked with SDHH. As such, teachers, and not students, are the unit of analysis in this study. Students are served in a wide range of educational settings, from separate schools for the deaf to fully inclusive regular education settings. In a break from previous years’ surveys, the Fourth National Survey moved from categories that describe the structure of the setting to the percent of time the students were in a separate setting or in a regular education setting. Previous surveys provided options such as “school for the deaf,” “regional program,” and “mainstreamed or inclusion program.” Participants had often indicated that more than one category applied to their students. In an effort to allow for more flexibility in responses, we created five categories for participants to use to describe their educational setting:

1. Separate setting for entire school day
2. Separate setting for half the day or more, but not the entire day
3. Separate setting for less than half of the day
4. Mainstreamed setting only
5. Other

These categories allow findings to be more closely aligned with other large-scale surveys of deaf education such as those conducted by the Gallaudet Research Institute and the Child Count for the IDEIA reports to Congress.

Recruitment

The Annual Survey followed the recruitment methods established in the previous 3 years of this research project, with a few exceptions. The first step was to contact all previous participants who had asked to be invited for future research efforts. Second, we contacted individuals who had requested that we notify them when new studies occurred, even if they had not previously participated in the project. Finally, we posted announcements on various online list serves and virtual meeting places for individuals who serve SDHH. Examples of these online resources include the mailing list for the Research on Education of Deaf Persons Special Interest Group (part of the American Education Research Association), Deafed.net, Self Help for the Hard of Hearing, and Pepnet.

Due to the potentially sensitive nature of the views shared in the study, participants were not required to provide identifying information to participate in the survey. Responses were therefore anonymous unless individuals chose to provide their contact information. Individuals who completed the survey were eligible to be entered into a prize drawing for $25 gift cards to Amazon.com, Target, or Starbucks. Persons who did not provide contact information in the survey itself were free to contact the principal investigator separately to enter the drawing, keeping their identifying information separate from the survey responses.

Survey Instrument

The Fourth Annual Survey was conducted both online (via www.surveymonkey.com) and via a paper-and-pencil format. A total of 231 (80%) participants used the online format and 59 (20%) participants used a paper-and-pencil format. There were three main sections to the survey: Demographics, Assessments and Accommodations, and Perspectives. The Demographics section of the survey included questions about the school name, district, and location (state) where the individual worked with SDHH. This information was gathered in order to monitor what areas of the country
had high representation and areas where we needed to focus our additional recruitment efforts. The Demographic section also asked participants to describe their role in working with SDHH and how many years of experience they had in those roles. The remainder of the Demographics section focused on classroom characteristics, such as the communication modes used in instruction, or on student characteristics, such as the number of SDHH and hearing students, cochlear implant use, and additional disabilities.

The second section of the survey focused on Assessments and Accommodations. The questions were primarily checkboxes of when SDHH used one of these five accommodations: student was given extra time for the test (ET), student took the test either individually or with a small group (SGI), test items were read aloud (orally) to the student (TIR), test directions interpreted (e.g., ASL or total communication) for the student (TDI), and test items interpreted (e.g., ASL or total communication) for the student (TII). This section also included a question about the extent to which accommodations were on an IEP but not implemented. The IEP implementation question had two components: a Likert scale (1 = never implemented to 5 = always implemented) for each accommodation and an open-ended comment box for participants to explain their ratings.

The Perspectives section of the survey had two questions. The first was an open-ended response item that asked participants to describe information they consider when evaluating the effectiveness of a test accommodation. These responses were coded using a thematic content analysis approach. Responses first were reviewed for clusters of similar themes across participants. The coding categories and a corresponding codebook were developed after discussions between the research assistant and the principal investigator. The research assistant then completed a full code of the data set using the categories. After the first full code, the research team discussed items that did not appear to fall into the coding scheme and revised the protocol to fit a larger proportion of the data set. The research assistant then did a second full code of the data set using the revised coding categories. The principal investigator reviewed all the codes and flagged items for further discussion. After this further conversation about the response, these items were coded either as “other” or in one of the coding categories. A summary of coding categories with examples is shown in Appendix A.

The second question was a set of Likert scale items related to issues that might affect the validity of an accommodated test score. These items were drawn from the qualitative results of the Third National Survey (Cawthon et al., 2009). The question began with the following definition of validity: Validity refers to the extent to which the accommodation reduces barriers to meaningful participation without changing the content of the test. A valid assessment is one that can be used to compare test scores with those students who did not have accommodations. In other words, is this test measuring what we think it is measuring? Participants then rated 10 items on the extent to which they consider the issue when choosing an accommodation for SDHH on statewide, standardized assessments. The survey used a Likert scale (1 = never considered to 5 = always considered) for this question.

Results

Demographics

The 290 participants in the Fourth Annual Survey came from 38 states and the District of Columbia. States with more than 10 participants included California (n = 19), Colorado (n = 11), Kentucky (n = 54), Massachusetts (n = 11), Minnesota (n = 16), South Carolina (n = 26), and Wisconsin (n = 13). These states typically had high responses rates from the state school for the deaf. Because these numbers represent teachers, and not schools or students, it is challenging to make a direct comparison to national population figures to know the representativeness of the sample. That said there is an overrepresentation of teachers from the southeast region of the country, a region that includes states such as Kentucky and South Carolina. As with the previous National Surveys, there is an underrepresentation from the northeast region of the country.

There was a great diversity in educational setting represented in this sample. A summary of survey participants by educational settings is shown in Table 2. Just over a quarter of participants (27%) worked with
SDHH in a separate setting for the entire school day. An additional 27% worked with students in a main-streamed setting. A much smaller percentage, 14%, had students who spend half of the day or more, but not the entire day, in a separate setting. Similarly 12% of participants had students who were in a separate setting for less than half of the day. The “other” category received a large number of responses, \( n = 53 \), or 18% of the sample. Examples of responses that fell into the “other” category included home schooling, on a case-by-case basis, or across multiple categories. These figures also show an overrepresentation of teachers working in schools for the deaf or other separate settings when compared with the national distribution of students and schools. Mitchell’s (2004) analysis of the IDEIA Count Data indicated that the Gallaudet National Survey also faced this sampling challenge. Mitchell estimates that 80% of students are receiving instruction in a regular education setting, with only about 20% in a separate school facility (such as a school for the deaf).

In addition to educational setting, we also collected data on the number of years teachers have worked with SDHH and the number of students they taught in 2006–2007. Overall, teachers had an average of 14 years of experience working with SDHH, with a range from 1 year to 40 years of experience.

Participants also served SDHH in different roles. Because participants could choose more than one category, there was an option for “multiple roles.” The majority (38%) of participants were teachers of the deaf; other roles included special education teachers (19%) and itinerant teachers (19%).

### Language Used in Instruction

Previous surveys indicated that the language use in instruction often played an important role in determining the kinds accommodations students received for assessment. Participants in the Fourth Annual Survey could indicate the use of one or more of the following forms of classroom communication: ASL, signed system such as cued speech or signed exact English, oral (speech) by the instructor, total communication or simultaneous communication by the instructor, or oral with a sign language interpreter. Participants could select more than one option if applicable. Indeed, there is a great deal of overlap between categories (i.e., the totals sum to greater than 100%). The most prevalent form of communication is some type of sign and speech together by a single individual (56%) followed by ASL (45%) and oral only (44%). These categories represent language systems in place when one person is providing the communication for classroom instruction. In contrast, only 25% of participants indicated the use of an interpreter in addition to speech by a classroom instructor.

### Student Characteristics

Although the teacher or educational professional was the main unit of analysis in this study, participants did provide some information about the students they served. First, the total number of students reported served across all participants was 1,083 hearing students and 5,903 SDHH. Because 23 participants did not provide student totals, this is an underestimate of the total student population served by these teachers. Of the SDHH taught by participants in this study, 581 were reported to wear a cochlear implant. Participants estimated the percentage of students at each of four levels of hearing loss: mild, moderate, severe, or profound. On average, \( M = 11\% \) (\( SD = 18\% \)) of a teacher’s students had a mild hearing loss, 20% (23) had a moderate hearing loss, 27% (24) had a severe

---

**Table 2** Participants by educational setting and years of experience

<table>
<thead>
<tr>
<th>Setting</th>
<th>Average number of years working with SDHH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separate setting entire school day (( n = 76 ))</td>
<td>13.8 (10.3)</td>
</tr>
<tr>
<td>Separate setting for half of the day or more (( n = 40 ))</td>
<td>13.0 (9.54)</td>
</tr>
<tr>
<td>Separate setting for less than half of the day (( n = 35 ))</td>
<td>15.3 (10.5)</td>
</tr>
<tr>
<td>Mainstreamed setting only (( n = 79 ))</td>
<td>15.1 (10.5)</td>
</tr>
<tr>
<td>Other (( n = 53 ))</td>
<td>14.4 (10.5)</td>
</tr>
<tr>
<td>Total (( N = 283 ))</td>
<td>14.3 (10.3)</td>
</tr>
</tbody>
</table>

Note. Only 283 of 290 participants provided information about the average number of years they worked with students who are deaf or hard of hearing (SDHH).
hearing loss, and 33% (31) had a profound hearing loss. Participants did not have information available for 9% of their students. Finally, we also collected information about additional disabilities of SDHH. A summary of the prevalence of different kinds of disabilities is in Figure 1. Figures are number of participants reporting at least one student with the additional disability. As with several items in this survey, participants could select more than one option. Only 27, or 9%, did not indicate another disability category. The most commonly selected additional disability was a learning disability, chosen by $n = 169$, or 60% of participants (recall that teacher is the unit of analysis in this figure).

Accommodations Use

The first research question addressed questions of accommodations prevalence. We collected data on the number of participants who had students use a range of test accommodations for mathematics, reading, and, where relevant, science. Teachers reported that 4,515, or roughly three-quarters of the SDHH served by this sample, participated in state standardized assessments using at least one accommodation. The survey included a checklist of five accommodations commonly used by SDHH. Participants indicated whether each accommodation was used for mathematics, reading, and/or science. A summary of accommodations use, by test subject, is shown in Table 3.

As expected, ET, SGI, and TDI were the most frequently used accommodations by teachers of SDHH. Prevalence rates on these accommodations ranged from 59% for TDI in reading to 74% for SGI, also in reading. They were also used fairly consistently across mathematics and reading, indicating that if a student receives an accommodation (such as TDI) in mathematics, it is likely he or she would receive it in reading as well. These are the most commonly allowed accommodations in state testing policies and are also seen as least threatening to the validity of the assessment.

TIR and TII are more controversial in terms of test validity because it changes the testing experience from one that includes a reading skill to one that uses an oral or visual presentation of the test item content. From past surveys, we have learned that TIR and TII accommodations tend to be less prevalent than ET, SGI, and TDI. The survey results in previous years have also shown some distinction in accommodations use between mathematics and reading, with higher levels of TIR and TII for mathematics than for reading. This basic finding was also borne out by this year’s results, though to a lesser extent than in previous surveys. TIR was checked off by 48% of survey participants for mathematics and by 36% of participants for reading. Results for TII were similar, with 49% and 33% prevalence for mathematics and reading, respectively. There is also some consistency “within” subject area here; TII and TIR results for mathematics are nearly identical to each other, and the same is true for reading.

This year was the first year the survey included science as a test subject option on the accommodations

<table>
<thead>
<tr>
<th>Accommodation</th>
<th>Mathematics</th>
<th>Reading</th>
<th>Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extended time (%)</td>
<td>65</td>
<td>67</td>
<td>48</td>
</tr>
<tr>
<td>Small group/individual</td>
<td>73</td>
<td>74</td>
<td>55</td>
</tr>
<tr>
<td>administration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test directions interpreted</td>
<td>60</td>
<td>59</td>
<td>46</td>
</tr>
<tr>
<td>Test items read aloud</td>
<td>48</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>Test items interpreted</td>
<td>49</td>
<td>33</td>
<td>37</td>
</tr>
</tbody>
</table>

Note. $N = 290$. Results reflect the percentage of teachers who had at least one student who is deaf or hard of hearing using the accommodation in each subject area.
checklist. Requirements under NCLB for testing science are different than for mathematics and reading in that science needs only to be tested in three grades, whereas mathematics and reading tests must be administered in third through eighth plus one high-school grade. As a result, there are fewer students taking science in any given year than mathematics or reading. Because this study uses the teacher as the unit of analysis, and not the student, it is difficult to distinguish between accommodations used for students in one grade compared to another. It is likely that fewer teachers have students taking science, and thus the study sample will report lower rates of accommodations use in that subject area. We therefore primarily looked at the accommodations for science results within the subject area. The patterns of accommodations use were actually quite similar in science as they were in mathematics and reading. For example, although ET results for science (48% of participants) were lower than those shown for reading (65%) and mathematics (67%), ET was still one of the more highly used accommodations in science. Other frequently used accommodations included TDI (46%) and SGI (55%), selected by approximately half of respondents. TIR (36%) and TII (37%) accommodations were the least prevalent of the five but were chosen by at least a third of the study participants. As with mathematics and reading, TIR and TII results were almost identical within science. Yet the figures are actually higher than one would expect given that in mathematics and reading, the percentage of teachers using TIR and TII was less than half of those using ET. In science, there was a four to three ratio in ET to TIR/TII use.

The above figures give an overall view of how participants selected accommodations use for mathematics, reading, and science. We also ran correlations to test the level of co-occurrence of individual accommodations across subject areas (Table 4). All correlations were significant at $p < .001$ but ranged in strength depending on the accommodation and the compared subject areas. For example in ET, correlations were $r = .86$, .65, and .69 for mathematics and reading, mathematics and science, and reading and science, respectively. Correlations for SGI were $r = .86$, .63, and .66 for the same relationships. For TII, results went in a different direction, with $r = .66$, .78, and .74 for mathematics and reading, mathematics and science, and reading and science, respectively. These results demonstrate that accommodations largely cluster together, with strong relationships in accommodation use across subject areas. As hinted at in Table 3, the correlations between science and the other two content areas is high for all, but even higher for test accommodations that involve either the use of sign language or having the test item read aloud (TDI, TII, and TIR).

### Effectiveness

Participants provided their perspectives on how they ascertained whether or not an accommodation was effective. These responses were coded and summarized across participants, as shown in Table 5. Responses to this question were quite varied and thus difficult to summarize into broader categories. The most commonly noted source of evidence was the students themselves, either their reactions during the test (18%) or whether they could understand the test (6%). If the purpose of the accommodation is to increase student access to test content, students are likely a good source of information about how well that process happens. The comparison with other test scores is another main theme that arose from this analysis. Options included comparisons between the accommodated score and class work (10%), nonaccommodated scores (1.7%), and scores of their peers (e.g., other students with disabilities, 4%). Other sources of evidence stemmed from the score itself, such as whether it was “high” (11%). For some individuals, a “high” test score meant that the accommodation made the test too easy. For others, a “high” test score meant that the accommodation was working as intended. The final large category reflect the idea that an IEP provides for valid accommodations, and that if

<table>
<thead>
<tr>
<th>Relationship</th>
<th>ET</th>
<th>SGI</th>
<th>TDI</th>
<th>TIR</th>
<th>TII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics–Reading</td>
<td>.86</td>
<td>.87</td>
<td>.91</td>
<td>.61</td>
<td>.66</td>
</tr>
<tr>
<td>Mathematics–Science</td>
<td>.65</td>
<td>.63</td>
<td>.74</td>
<td>.73</td>
<td>.78</td>
</tr>
<tr>
<td>Science–Reading</td>
<td>.69</td>
<td>.66</td>
<td>.76</td>
<td>.66</td>
<td>.74</td>
</tr>
<tr>
<td><strong>Note.</strong> All correlations are significant at $p &lt; .001$. ET = extended time; SGI = small group or individual administration; TDI = test directions interpreted; TIR = test items read aloud; TII = test items interpreted.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the accommodated test score aligns with those stipulations, that it was an effectively accommodated test (10%). Example quotes for each category are provided in Appendix A.

Issues That Impact Validity

This last section built upon previous qualitative work that identified key factors that might have an impact on the validity of an accommodated test score (Cawthon, Wurtz, & the Online Research Lab, in press). When asked about score validity, educators who worked with SDHH identified student characteristics, classroom practices, and systemic factors that might increase or decrease the accuracy of the assessment. Participants in the current study rated 10 factors on the extent to which they considered its effect on validity when choosing accommodations for a statewide, standardized assessment. Participants rated each item on a scale of one to five, with 1 = never considered and 5 = always considered. Results are shown in Figure 2. All the factors listed were rated highly; the lowest item, Student Views, still had a mean score of 3.52.

<table>
<thead>
<tr>
<th>Table 5 Evidence for effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response category</td>
</tr>
<tr>
<td>Student reaction during test</td>
</tr>
<tr>
<td>Consistency between class work and test grades</td>
</tr>
<tr>
<td>Whether or not the test score is high</td>
</tr>
<tr>
<td>Whether or not the student requested the accommodation</td>
</tr>
<tr>
<td>Student can understand test directions and items</td>
</tr>
<tr>
<td>Comparing scores with other students with disabilities</td>
</tr>
<tr>
<td>Looking at test content</td>
</tr>
<tr>
<td>Individualized Education</td>
</tr>
<tr>
<td>Program versus what was used on the assessment</td>
</tr>
<tr>
<td>Accommodated versus not accommodated test scores</td>
</tr>
<tr>
<td>Whether there was an unfair advantage</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>

Note. Participant responses could be coded for one or more categories. A total of 194 participants (67%) provided responses. Figures represent the number of codes that were assigned to that topic out of a total of 300 codes assigned across all participant responses.

Discussion

The results of this study add richness to the available literature on important issues in accommodations and standardized assessment for SDHH. Before addressing the research questions that guided this study, the article first articulates important limitations to the study design that must be considered when drawing conclusions about the findings presented here.

Limitations

The first limitation to this study is the unit of analysis: individual teachers or educational professionals that work with SDHH. For the research questions regarding the prevalence of accommodations use in mathematics, reading, and science, the unit of analysis makes it difficult to draw a direct comparison between subject areas. This is true because fewer grades, and thus fewer students, are required to participate in science assessments. When teachers report the prevalence of accommodations use as “at least one of their students” and not by individual student, it is difficult to interpret findings about the relative use of accommodations for science versus other subject areas that are required for more students. As a further note, these teachers are largely serving students at either schools for the deaf or programs for the deaf or hard of hearing. They are thus serving larger groups of students than a teacher in a mainstreamed setting who might only have one or two students. This limitation means that the results here are only a rough estimate of practices and reflect the perspectives of teachers who are in a setting that is
focused more specifically on the pedagogical needs of SDHH.

A second limitation is the lag time between the time that students receive accommodations and the time that teachers participated in the Fourth Annual Survey. In general, the survey is designed to ask teachers to report on the previous assessment year (in this case, 2006–2007) during the subsequent academic year (for this study, 2007–2008). The year time lag is necessary in order to capture the assessment process for an entire cohort of students (up until the very end of a school year) and still be able to contact teachers over a semester’s time frame (which occurs after the summer and into the fall of the following school year). A related issue inherent in accommodations decision making is the low reliability of recommendations, even within the same individual for the same case. It is reasonable to think that the reliability of reported information in this survey may also be low, with low internal consistency of reported prevalence or views on accommodations (Cawthon, 2009; Tindal, 2007). The findings here represent a snapshot at one point in time and may not accurately reflect views the teachers had during the assessment process as it had occurred the year prior.

The third limitation applies mainly to the findings about teacher perspectives regarding how they evaluated the effectiveness of an accommodation. Only a subset of the participants offered justifications for their ratings about accommodations implementation \( n = 194 \), 67% of the overall study sample. The responses here may not be representative of the other participants who did not answer this question. It is also possible that participants did not know or had not considered this issue before and thus left the question box blank. Findings from this study should be considered with these limitations in mind.

Test Accommodations in Mathematics, Reading, and Science

The first research question guiding this study was a focus on accommodations use for the three required content areas under NCLB: mathematics, reading, and science. Results for mathematics and reading were as expected given results from the previous research and confirm the high use of accommodations for SDHH of teachers who participated in this study. The three most frequently used accommodations, ET, SGI, and TDI, are also the three that are most commonly allowed by state assessment policies (Clapper et al., 2005). Their use is largely thought not to pose a threat to the validity of the resulting test score and thus can be used by students participating in high-stakes assessment (Sireci et al., 2005). Accommodations that may pose a threat to validity of the test score (TII and TIR) are used less often for SDHH than ET, SGI, and TDI. For example, results here indicated a moderate correlation between TII use in mathematics and reading \( r = .66 \), a figure that is lower than the correlation for ET \( r = .86 \), SGI \( r = .87 \), and TDI \( r = .91 \) use in these two subject areas. Furthermore, TII and TIR are used the least often for reading assessments, the area of greatest concern from a test validity perspective. The consistency of these findings for mathematics and reading across assessment years moderates concerns about the reliability of teacher reports within each year of the survey.

New to the NCLB assessment system, and also to this analysis, was the inclusion of science as a subject area. Although the overall prevalence of accommodations was lower for teachers of SDHH who took science, considering the limitations of teachers as the unit of analysis in this study, what is most important is the “relative” use of accommodations within this subject area as compared with mathematics and reading. It does appear that the three “high-use” accommodations for mathematics and reading (ET, SGI, and TDI) are also the most frequently used in science (48%, 55%, and 46%, respectively). The correlation tables for these three “high-use” accommodations indicate a medium to high relationship between their use in science with the other areas. For example, the correlation of ET use in science and mathematics was \( r = .65 \) and for science and reading it was \( r = .86 \).

Results indicate a different trend for accommodations that change the presentation of the test item. We would expect a lower overall frequency for TII and TIR because of the policy limitations and the fewer numbers of students who might need these accommodations. This is certainly the case, with overall reported use by 36% (TII) and 37% (TIR) of participants. When compared with the other subject areas, we see some differences in their use. Teacher reports
of TII and TIR accommodations use in science seem to be similar to their use in mathematics. Using ET as an example, the ratio of ET for mathematics to science (65%/48%) is 1.35. If TII and TIR were being used in the same relative frequency for mathematics and science as ET, one would expect a similar use ratio (i.e., no proportional drop). We do see this TII and TIR, with ratios for mathematics to science of 1.32 and 1.33, respectively. For reading, however, we see a different pattern. The ratio of reading to science is 0.89 for TII and 1.00 for TIR. These ratios are unexpected given the ratios seen for other accommodations. TII and TIR use in science is thus higher than one would expect if teachers were using the accommodations consistently across all subject areas. This result fits with the nature of science content and assessment format discussed in the literature review in this article.

An important area of future research is to investigate how TII and TIR are used with different kinds of test items within each subject area. In other words, analysis of accommodations use and effectiveness is needed at the item level, or at the very least the domain level on an assessment. For example, it would be relevant to know if SDHH use TII and TIR accommodations on test items that require students to read a passage of text, a poem, or other block of written narrative. Depending on the task demand, TII or TIR accommodations may change what the student needs to do to accurately respond to the test item. On the other hand, if TII and TIR are used more frequently on test items that require the student to show their knowledge through a student-constructed essay, mathematics graph, science concept map, or lab demonstration, there may be appropriate and useful uses of these accommodations. In a sense, for student-constructed response items, TII and TIR on test items may serve the role of providing instructions, such as when one has TDI, than in providing information about the test question and its relationship to the response choices.

Evidence of Effectiveness

What evidence teachers of SDHH used to determine effectiveness is also an emerging area of research in this field. In essence, this question focuses on the “epistemology” of the teachers in order to better understand the reported practice of accommodations use with their students. This question was exploratory in nature; our understanding of assessment accommodations epistemology relied on answers to an open-ended question on how participants determined the effectiveness of accommodations for their students. The question assumed that teachers do evaluate effectiveness of accommodations as part of their practice, which may or may not have been the case for all participants.

The first finding that is salient here is the wide range of responses to this question, with consensus only on a few key categories (see below). This result suggests that the participants in this study sought evidence of effectiveness from very different sources. For example, some teachers focused on compliance (e.g., if the accommodation was on the IEP), whereas others looked at the test content to decide of the accommodation was appropriately assigned. Although most participants were teachers for the deaf, and thus had a similar training background, their classroom experiences and individual needs of their SDHH may lend to this great diversity of perspective. When considering patterns of decision making among educational professionals who serve SDHH, this diversity of epistemologies of “how do we know if accommodations work” is important to keep in mind.

When looking at overall patterns of epistemology shown in this analysis, teachers largely focused on student satisfaction in their evaluation of accommodations effectiveness. If a student could use the accommodation, was not frustrated with the test process, and could understand what was being asked of him on the assessment, then the accommodation was seen as effective. In a related vein, some teachers noted that if the student took the initiative and requested the accommodation, then it must have been effective. This is particularly true for students in middle or high schools who often can only have the accommodation if they request it themselves. Student affect and ease of participation in the assessment are critical components of this student satisfaction epistemological perspective.

A third interesting finding from this analysis is that some teachers also looked for convergent validity from other sources beyond the accommodated test score. Teachers looked for consistency between the
accommodated test score and student performance on classroom assessments or other students with disabilities. In both cases, the accommodation was seen as effective if student performance on the standardized assessment meets expectations as set by other sources. The relative validity of the other sources was not mentioned but presumed to be an anchor for comparison with the accommodated standardized assessment. In a related vein, and using a “score boost” perspective used by researchers who investigate the effectiveness of accommodations, a few teachers noted comparing test scores for the student when they used the accommodation and when they did not. Finally, a group of teachers said that they considered the absolute nature of the score, whether or not it was “high.” Interestingly, a “high” score was construed as evidence of either a good accommodation or a bad accommodation, depending on the teacher. Overall, these approaches reflect a test score validity epistemology of accommodations effectiveness.

In this study, we looked at issues that affect validity as a second way to understand how teachers think about accommodations and their effectiveness. In a sense, this question returns to the concept of effect of accommodations and whether or not it fairly changes test scores. Recall the statement participants responded to on the survey: Validity refers to the extent to which the accommodation reduces barriers to meaningful participation without changing the content of the test. A valid assessment is one that can be used to compare test scores with those students who did not have accommodations. In other words, is this test measuring what we think it is measuring? Participants then rated items on the extent to which they consider the item and its effect on validity when choosing an accommodation for SDHH on statewide, standardized assessments. Whereas previous research asked teachers to brainstorm what issues are important when making decisions about accommodations for SDHH, this study looked specifically at 10 important components and explicitly tied the question to the issue of validity.

All 10 of the items listed on the survey received mean ratings of at least 3.5 (out of a scale of one to five). This is perhaps not surprising due to the fact that the factors themselves were drawn from previous responses of teachers of SDHH, individuals who are similar in background and professional roles as this sample. Issues such as student language proficiency (both in English and in ASL) and whether ASL was used in instruction are especially salient for the SDHH population. The highest items were reading proficiency \((M = 4.37)\) and opportunity to learn \((M = 4.35)\), followed by content area proficiency \((M = 4.16)\) and if instruction included use of a sign system of other than ASL \((M = 4.1)\). One thing these data illustrate is that these factors are not just generally of interest to educators who serve SDHH but that they are a specific part of how professionals think about what goes into a valid accommodated test score for SDHH. A second finding in these data is the great complexity of the accommodations decision-making process for IEP teams and teachers of SDHH. When a student has multiple languages in their home and school environment as well as a disability that prevents access to some elements of classroom instruction, accommodations play an important role in facilitating the assessment process. Other factors not studied here may include a teacher’s own preparation in the content area and how that might influence the accommodations decision process. How all these elements are organized when making decisions about accommodations for SDHH, and what weight each may play when there are competing needs, is largely unknown.

**Conclusions**

Standardized assessment for accountability purposes has become a formalized component of how educational policy works to enhance educational outcomes for all students. For SDHH, meaningful participation in standardized assessment often requires the use of accommodations that increase student access to test content. Larger questions regarding the prevalence accommodations use have seen some stabilization in recent years, with teachers and IEP teams largely utilizing accommodations resources available for their SDHH. Areas still in need of investigation include measurement of the effects of test accommodations on student performance on assessments and evaluation of the resultant psychometrics of accommodated assessments. This need is true for mathematics and reading, but as discussed here, takes on an even more complex implication when looking at the task demands in science as a content area.
The purpose of this study was to look at two emerging areas that affect accommodations practice: science as a content area and how teachers evaluate the effectiveness of accommodations. In a sense, both these issues relate to an epistemology about assessment for SDHH. In science, students are asked to demonstrate that they gather evidence and use scientific inquiry in ways that reflect their development as users and consumers of science. Accommodations for science thus focus not just on access to the test item but what will facilitate the demonstration of knowledge in student-constructed response items. In assessment decision making, teachers are asked to evaluate and gather evidence regarding the match between the student, the test content area, and the chosen accommodations. Emphasis on one area over another, such as the student satisfaction and test validity perspectives found in this study, might prove explanatory as to why decision-making process is less systematic than one might expect. For both emerging issues of epistemology, what constitutes sound evidence for SDHH reflects the diversity of students and those who serve them.

**Funding**

The author was partially supported by a University of Texas at Austin special research grant (2008–2009).

**Appendix A**

**Table A1**  Example responses to evidence for effectiveness

<table>
<thead>
<tr>
<th>Response category</th>
<th>Example response in this category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student reaction during test</td>
<td>I consider whether the student will use the accommodation or whether he/she will see it as a negative. If the student feels successful on the test with the accommodations, then the accommodation was effective.</td>
</tr>
<tr>
<td>Consistency between class work and test grades</td>
<td>We compare our classroom-based knowledge of the student’s performance with the results of the interpreted mathematics statewide test.</td>
</tr>
<tr>
<td>Whether or not the test score is high or higher</td>
<td>Scores were increased and some students passed the exit exams.</td>
</tr>
</tbody>
</table>

**Table A1** Continued  Example response in this category

<table>
<thead>
<tr>
<th>Response category</th>
<th>Example response in this category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whether or not the student requested the accommodation</td>
<td>If the student initiated its use, was able to participate in the test with the accommodation</td>
</tr>
<tr>
<td>Student can understand test directions and items</td>
<td>If the student “understood” what was asked/what to do, then did it</td>
</tr>
<tr>
<td>Comparing scores with other students with disabilities</td>
<td>Look at the average of deaf students’ scores and compare</td>
</tr>
<tr>
<td>Looking at test content</td>
<td>If a student is far behind grade level, especially in reading and language, I will read aloud or sign test questions in science, history, and mathematics for two reasons. First, I do not want the student to be too frustrated. Second, I want to test the content knowledge, not just the reading comprehension on these tests.</td>
</tr>
</tbody>
</table>

**Individualized Education Program versus what was used on the assessment**

Our students must have the same accommodations daily in order for them to be used on statewide testing. We are able to adjust accommodations as needed on the Individual Education Plans prior to the actual testing situation.

**Accommodated versus not accommodated test scores**

We also compare how the student had done the previous year without interpretation to this year with the interpretation.

**Whether there was an unfair advantage**

If the curriculum has been consistently presented with accommodations throughout the year, and the students know the material, the goal of accommodations is to make it possible for them to show it … not to make anything easier or “better” than their hearing peers.

**Other**

A wide range of topics with a single response
Notes

1. “Teachers,” “Educators,” and “Participants” are used interchangeably in this article.
2. “Validity” is a term with multiple meanings. In this context, validity refers to the accuracy of the interpretation of the test score. If the interpretation is valid, then the user (teacher, accountability reform, etc.) can make decisions that are based on an accurate representation of a student’s knowledge or skill. In this discussion regarding accommodations, validity may change because the accommodation increases or decreases the accuracy of the test score.

References


Received September 29, 2009; revisions received January 6, 2010; accepted January 7, 2010.