Computer-Based Signing Accommodations: Comparing a Recorded Human with an Avatar[1]

Michael Russell
Maureen Kavanaugh
Lynch School of Education, Boston College

Jessica Masters
Jennifer Higgins
Technology and Assessment Study Collaborative, Boston College

Thomas Hoffmann
Nimble Assessment Systems

Abstract

Many students who are deaf or hard-of-hearing are eligible for a signing accommodation for state and other standardized tests. The signing accommodation, however, presents several challenges for testing programs that attempt to administer tests under standardized conditions. One potential solution for many of these challenges is the use of computer-based test delivery that integrates recordings of signed presentation of test content into the test. In addition to standardizing conditions, computer-based delivery holds potential to decrease the cost of developing recordings of signed presentation by using avatars rather than humans. However, because avatars are relatively new and are not as expressive or lifelike as humans, they may not be as affective as humans in presenting content in a clear and interpretable manner. The study presented here employed a randomized trial to compare the effect that a computer-based provision of the signed accommodation using a recorded human versus a signing avatar had on students' attitudes about performing a mathematics test and on their actual test performance. This study found that students generally reported that it was easy to perform a mathematics test on computer, and that both the recorded human and the signing avatar tools were easy to use and to understand. Students also reported a strong preference for performing future tests on computer, and generally preferred using the recorded human and the avatar for future tests rather than a DVD. While students also reported that they preferred the recorded human rather than the signing avatar, this preference did not affect test performance. The use of the recorded human and the avatar did not have affects on either the amount of time required to complete the test items or on students' performance on the test items. Implications for future research are discussed in light of these findings and the shortcomings of this study.

Keywords: Computer-based Testing, Test Accommodations, Deaf and Hard-of-Hearing

[1] The computer-based test delivery software and the methods used to provide a signing accommodation on computer used for this study were developed by Russell and Hoffmann. To reduce the appearance of a conflict of interest, statistical analyses were conducted by Jessica Masters, who is unaffiliated with the development of the software and methods.

Computer-Based Signing Accommodations:
Comparing a Recorded Human with an Avatar

Many students who are deaf or hard-of-hearing are eligible for a signing accommodation for state and other standardized tests. The signing accommodation is intended to assist students who are deaf or hard-of-hearing, and whose reading skills are below grade level, access test content. As is described more fully below, the signing accommodation presents several challenges for testing programs that attempt to administer tests under standardized conditions. One potential solution for many of these challenges is the use of computer-based test delivery that integrates recordings of signed presentation of test content into the test. In addition to standardizing conditions, computer-based delivery holds potential to decrease the cost of developing recordings of signed presentation by using avatars rather than humans. However, because avatars are relatively new and are not as expressive or lifelike as humans, they may not be as affective as humans in presenting content in a clear and interpretable manner. The study presented here is a first attempt to compare the effect of the use of recorded humans versus an avatar to present a signed accommodation for a mathematics test has on students' attitudes, preferences, and actual test performance.

The article presents findings from a pilot study that focused specifically on providing a signing accommodation on computer using a recorded human interpreter or an avatar, and compares the effects of both methods on student preferences for how a singing accommodation is delivered and on test scores. The article begins by examining current shortcomings to the provision of signing accommodations for students who are deaf or hard-of-hearing. We then discuss the potential benefits that computer-based administration may have for standardizing and individualizing the provision of a signed accommodation, and the role that avatars may play for providing a computer-based signed accommodation. The method used to compare a recorded human with an avatar is then presented along with the results and potential implications of the findings.

Background

Prior to the 1990s, students with disabilities were generally excluded from large-scale assessments. In an effort to assure that students with disabilities and special needs receive appropriate instruction, the federal *No Child Left Behind Act of 2001* (NCLB, 2002) and *Individual Disability Education Act* (IDEA, 2004) require students with disabilities and special needs to participate in statewide assessment programs. While these recent legislative efforts have been successful in increasing participation of students with disabilities in statewide assessments, successful inclusion requires states to provide students with appropriate opportunities to demonstrate their abilities. Test accommodations are the most common method for meeting this requirement.

Test accommodations are defined as a change in the way a test is administered or how a student responds to a test item and are "intended to offset distortions in test scores caused by a disability without invalidating or changing the test measures" (Elliot, Kratochwill, & Schulte, 1999, p. 2). Test accommodations are intended to allow students with disabilities the opportunity to demonstrate knowledge and participate in statewide assessments on a more equal playing field as their non-disabled peers (Driscoll, 2007) and should be recommended based on

the needs of each individual student (Horvath, Kampfer-Bohach & Kearns, 2005). While state policies vary considerably regarding which types of accommodations are allowed and how decisions are made regarding which students may be provided with a given accommodation (Thompson & Thurlow, 2001), one of the most common accommodations provided is reading aloud test content (Bielinski, Thurlow, Ysseldyke, Freidebach, & Freidebach, 2001).

Depending on the level of reading support required by the student, the read-aloud accommodation consists of reading aloud only the test directions or the directions and the test items. The argument for providing read-aloud accommodations is that the accommodation removes irrelevant barriers (e.g., poor reading ability, reading disability or coding difficulties) to the construct being measured (i.e., content area knowledge) and helps students with reading disabilities and ELL students access the test material (Bielinski, Thurlow, Ysseldyke, Freidebach, & Freidebach, M., 2001; Abedi, Hofstetter, Baker, & Lord, 2001; Tindal, Heath, Hollenbeck, Almond, & Harniss, 1998). This argument, however, only applies to tests that are measuring constructs that are distinct from the ability of students to decode text or to decode and comprehend text. As summarized by Sireci, Li, & Scarpati (2003), a substantial body of research indicates that, for students who have difficulty reading test content, the read-aloud accommodation is generally effective for improving scores on tests that measure constructs that are unrelated to decoding text.

For students who are deaf or hard-of-hearing, the read-aloud accommodation is provided in either American Sign Language or Signed English. While ASL and Signed English both employ hand gestures and symbols to represent words, phrases, and expressions, the two forms of communication are distinct from each other. American Sign Language is a unique form of communication that has its own set of rules for grammar and often employs word order that differ noticeably from spoken English. Signed English employs the same grammatical rules as spoken English and effectively represents words through hand gestures. Although most people who communicate in ASL can also communicate in Signed English, people who communicate in Signed English often are not familiar with ASL. Nonetheless, since ASL and Signed English require translation of text into an alternate form of communication, this form of the read-aloud accommodation is generally referred to as a "signing accommodation." According to Lazarus et al. (2006), 45 states allow signing accommodations for students who are deaf or hard-of-hearing. A recent survey of 258 schools and programs serving students with hearing impairments found the use of a sign language interpreter for directions was the second most frequently used accommodation, with only extended time used more frequently for this group of students (Cawthon, 2006). Fifty-seven percent of respondents also reported that an interpreter was used to interpret items for mathematics assessments and 41% for reading assessments.

Shortcomings of the Signing Accommodation

The signing accommodation typically requires the use of an "access assistant" (i.e. readers, sign language interpreters) who provides students with specialized support to improve accessibility to test content (Clapper, Morse, Lazarus, Thompson & Thurlow, 2005a). As with scribes and readers, sign language interpreters require extensive training and multiple checks by other experts to ensure consistency in the interpretation of test materials, which is critical in the provision of testing accommodations (Clapper, Morse, Thompson & Thurlow, 2005a). The use of access assistants who have not received proper training decreases the quality of the accommodation and interferes with the accuracy of test scores. Moreover, the use of assistants who vary in their signing ability and in their familiarity with the test content results in the

inequitable provision of the signing accommodation to students across a testing program. This problem is particularly concerning given the shortage of trained and qualified sign language interpreters, making it more likely that some students will be provided assistants with lesser experience and qualifications (Johnson, Kimball & Brown, 2001). In turn, the variability in proficiency of the signing assistants is likely to have a negative effect on test validity and the comparability of test scores for students who use an access assistant (Clapper, Morse, Thompson & Thurlow, 2005b).

Johnson, Kimball and Brown (2001) found that sign language interpretation as a testing accommodation poses several practical and psychometric challenges. They cited lack of training as major issue in their observational study of sign language interpretation of large scale mathematics and reading assessments. Participating interpreters were sometimes unfamiliar with or misunderstood the specific state guidelines for appropriate use of accommodations. Many were also reported to be unfamiliar with mathematics vocabulary, especially at the higher grades, which in turn affects the quality of interpretation. The importance of maintaining test security has made overcoming this obstacle increasingly difficult as interpreters are not permitted access to test materials prior to the day of testing. Many interpreters reported feeling unprepared to interpret lengthy passages or unfamiliar mathematics terms. Teachers also pointed out practical issues that pose problems for students who receive this type accommodation. For example, sign language interpretation requires the student to look up at the interpreter, making it difficult for them to make notes about specific aspects of test item simultaneously (e.g., to highlight key words or represent a narrative description of a problem as a mathematical equation).

The lack of clear and uniform state guidelines contributes to the inequitable provision of high-quality signing accommodations. Clapper, Morse, Thompson and Thurlow (2005b) found that only a handful of states provide clear and specific guidelines for scribes, readers and sign language interpreters, the three most frequently used access assistants. They also found that of the 22 states that provided guidelines for one or more type of assistant, only 11 had guidelines for sign language interpreters and 6 had guidelines that specified qualifications or characteristics of sign language interpreters. The breadth and depth of these guidelines also varied from state to state (Clapper, Morse, Thompson & Thurlow, 2005a; Clapper, Morse, Lazarus, Thompson & Thurlow, 2005b).

Providing Standardized Signing Accommodations Via VHS and DVD Recordings

One strategy employed by a few states that overcomes the problems presented by the use of signing assistants is the provision of the signing accommodation in a video-taped format. For this form of the accommodation, a single signer is used to present the test content to all students. Prior to producing a video-taped recording of the signer, the testing program works with the signer and other experts familiar with the signing of the test content to assure that the material is presented accurately and clearly. A single recording is then made and distributed to all students requiring a signing accommodation. In turn, the provision of high-quality signing is standardized across the entire testing program.

While standardization is highly desirable for a standardized testing program, the use of video-taped recordings are unsatisfactory for four reasons. First, it does not remove the physical separation between the test content and the presentation of that content. In most cases, the recorded accommodation is played on a television screen that sits in front of the classroom while the student works on a test booklet on his desk. This physical separation makes it difficult for

students to take notes or to move between the test item presented on paper and the signing of the text. Second, the use of a VHS or DVD machine to play the signed version requires a teacher or the student to use a remote control device to reverse or fast forward through the tape in order to replay specific portions of the tape. This inefficiency increases the time of testing and causes the focus to move away from the tested content onto the use of the controller. Third, in some cases, the video is played to a group of students simultaneously, requiring individual students to make a request in front of their peers to have a portion of video replayed and to then force all students in the room to re-watch that section. Fourth, when played for a group of students, the distance between the video and the student can make it difficult for students to clearly view the signed delivery.

Computer-based Delivery of the Signing Accommodation

Computer-based delivery of a test, however, holds promise to overcome problems associated with providing the signing accommodation either by a human or through a VHS or DVD machine displayed on a television screen. By embedding signed video directly in a computer-based test, the video and the test content can be presented in a consistent manner to all students and the signed representation can be thoroughly reviewed and approved prior to the day of testing. Embedding signed video in a computer-based test also allows the examinee to view a signed presentation of test content in very close proximity to each other and to the student (e.g., the video can be displayed within inches of the text-based representation of the item). Students can also be provided with individualized control over the size of the video displayed on their computer screen. Segments of video can also be linked to blocks of text or portions of an item (e.g., each answer option) such that a student can click on the text and the associated video is played automatically, thus eliminating the need to use a controller to fast forward or reverse through video. Finally, because the video is played on each individual computer, students may view portions of a video as many times as needed without affecting the test experience of other students.

Computer-based delivery of signed accommodations also provides an opportunity to employ signing avatars to present the signed accommodation. Avatars are human-like digital figures that can be programmed to move like a human. Avatars are being used increasingly for training videos, particularly by the military, because the cost of producing video with an avatar is significantly less than producing video using live human actors. Given the high cost of producing signed video, savings that may result through the use of an avatar are attractive.

In addition, the use of an avatar provides opportunities for students to customize specific aspects of the signed accommodation. Avatars base their movements on a script. A single script can be used to make multiple avatar characters move in exactly the same manner. This, then, allows students to select the character they like best from among a set of characters. Thus, a student could select a character that is male or female, short or tall, brown haired or blond haired, dark skinned or light skinned, etc. Giving students the choice of character may help increase their connection with the character and in turn their motivation during testing. Similarly, students could also control the background color in front of which the avatar signs test content. Controlling the background color would allow students to optimize the contrast between the avatar and the background so that they can best view signed presentation of the test content.

In addition, students can also be given control over specific features of the signed presentation. For example, a student who reads lips while communicating via Signed English could activate lipping, which allows the avatars lips to move in manner that imitates the speaking

of words. Similarly, a student with partial hearing who both views Signed English and obtains verbal cues could activate sound that works in conjunction with the signing. Finally, a student who is equally fluent in ASL and Signed English could switch between the two at any time during testing. Each of these features of avatars may help students tailor the provision of signing accommodations to meet their specific individualized needs. Table 1 provides a summary of the pontetial advantages and disadvantages of providing a signed accommodations via a human, video-tape player, recorded human embedded in a computer-based test, and an avatar embedded in a computer-based test.

Table 1: Advantages and Disadvantages of Signing Accommodation Provided by a Human, Video Player, Embedded Computer-based Test, and a Computer-based Avatar

Method	**Advantages**	**Disadvantages**
Human	• Natural body movements • Students accustomed to communicating with humans	• Difficult to standardize across settings • Misrepresentation of content • Unequal access to quality signing across settings • Often creates large physical separation between the test content and signed presentation
Video Player	• Controlled, high-quality presentation by a qualified interpreter • Equitable across settings	• Requires a state testing program more time and cost to produce and distribute than human • Typically presented in group settings • Creates large physical separation between content and signed presentation • Requires use of a video player control device
Video of Human Embedded in Computer-delivered Test	• Controlled, high-quality presentation by a qualified interpreter • Equitable across settings • Close proximity of content to signed interpretation • Does not require use of video control devices • Presented individually to student • Allows student to easily re-watch portions of the signed presentations • Able to track the use of signing for each individual item • Allows student to alter the size of the video window	• Requires a state testing program more time and cost to produce and distribute than human • Requires each student to have access to a computer • Requires students to be familiar with using a computer
Avatar Embedded in	• Same as above plus • Can allow student to select avatar	• Same as above plus • Avatar body movement is not as

Computer-delivered Test	• Can allow student to include or exclude captioning • Can allow student to activate or deactivate lipping • Can allow student to turn on sound that accompanies the signed presentation (for partial hearing) • Less expensive than recorded humans • Can allow student to switch between ASL and Signed English	smooth as a human • Students may be unfamiliar with avatar

Despite these advantages, some concern has been raised by members of the deaf and hard-of-hearing community about the use of avatars to provide a signing accommodation for tests. As an example, when the idea was discussed for the Massachusetts state test, some teachers of students who communicate in ASL raised concerns that unfamiliarity with avatars and the fact that the avatar's movements were not as lifelike as a human may have a negative effect on students test performance (Dan Weiner, January 26, 2008 personal communication). Since the lack of familiarity with avatars is not unique to students who communicate in ASL, these same concerns apply to students who communicate in Signed English.

The development of technology that enables the creation of signed avatars is relatively new. For this reason, very little research has focused on the use of signing avatars in an educational setting. In fact, a review of the literature found no peer-reviewed published articles that examined the use of signing avatars. However, one unpublished study conducted by researchers at Center for Applied Special Technology (CAST) explored the use of a signing avatar to develop students' reading skills. This study examined the use of signing avatars embedded into a product named Thinking Reader that is designed to help students develop reading comprehension strategies and skills. The study found that after a short period of time, students became comfortable working with the signing avatars, that they would often watch avatars sign material multiple times, and that the reading skills of students increased through the use of Thinking Reader with embedded signing supports (Dalton, Kennedy, Lutz, & Schleper, 2003).

Given the lack of prior research on signing avatars, concerns that students may be unfamiliar with signing avatars, and concerns that students may find an avatar's movements offsetting, before avatars are used for computer-based tests, it is important to examine the extent to which their use affects students' attitudes about taking tests and their actual test performance. The study presented here is a first effort to compare the use of a signing accommodation provided by human video versus an avatar. It is important to note that, while there are several issues related to signing accommodations that are in need of further research, this pilot study focused narrowly on the attitudes of students and effect of embedding a video-taped human or an avatar into a computer-based test. Findings from this study are intended to inform decisions about the use of signing avatars when providing a signing accommodation and to inform the design of future studies that empirically examine differences in preference and performance among a human, video-taped, and embedded computer-based signing accommodation.

Research Methodology

The study presented here aimed to compare the use of recorded human signing versus an avatar for a computer-based mathematics test. For this study, we were interested in: a) the extent to which the use of a signing tool that was integrated into a computer-based test was intuitive; b) students' attitudes about taking a test on computer using the human video and the avatar; and c) the effect that the use of an avatar had on student performance compared to the use of a recorded human. Specifically, we were interested in whether students preferred one method over the other, and whether the use of the two methods of presenting a signed accommodation affected students' test performance. Given the potential cost savings associated with the avatar and the potential ability for students to use an avatar to customize the signing accommodation, we analyzed whether any difference in preference translated into differences in test performance.

For this study, a convenience sample of 96 students in middle and high school who communicated in American Sign Language and attended schools for students in Connecticut and Massachusetts that served students who were deaf or hard-of-hearing performed two sets of 10 mathematics items. The students ranged from grade 8 to grade 12, with the vast majority being in grades 9, 10 and 11. All students were considered "fluent" in communicating in American Sign Language and used ASL to communicate with teachers and peers in the classroom. The two participating schools served only students who were deaf and hard-of-hearing, so students were in classrooms with only other students who communicated in American Sign Language. Due to IRB limitations, data on student prior achievement in mathematics and other demographic information, including each individual's specific grade level, were not collected.

The mathematics items were released Grade 8 National Assessment of Educational Progress items. Anticipating that some students may be performing below grade level in mathematics, only items with moderate to easy item difficulty were selected. To ensure that the items did not measure content to which most participating students had not been exposed, the items were shared with teachers in the school, who informed us that the items were appropriate for the vast majority of students participating in the study.

The two test forms were matched by item difficulty such that the mean item difficulty (based on the national sample of students performing each item) was .622 for Form 1 and .623 for Form 2. Items ranged in difficulty from .47 to .89, with 13 items within the range of .54 to .66. Students were randomly assigned to perform the first 10 items using either the recorded human or the avatar. The second ten items were then presented using the alternate method.

All test items were presented on computer using NimbleTools. NimbleTools is a federally funded computer-based test delivery system that was developed using principles of universal design to flexibly tailor the test delivery interface to meet the accessibility needs of each individual student (Russell, Higgins, & Hoffmann, 2009). NimbleTools includes 18 different accessibility tools that include various methods for reading aloud text, tab-entering for navigating test content and the testing interface, multiple magnification tools, multiple color contrast tools, an auditory calming tool, and a signing tool. For this study, only the signing tool was activated. Moreover, as described above, the signing tool was programmed to present the signing accommodation in American Sign Language using either a recorded human or the avatar.

Standardizing the Human and Avatar Signing

For this study, the signed presentation of content was developed in a manner that was as similar as possible when either the recorded human or the avatar was used. To accomplish this,

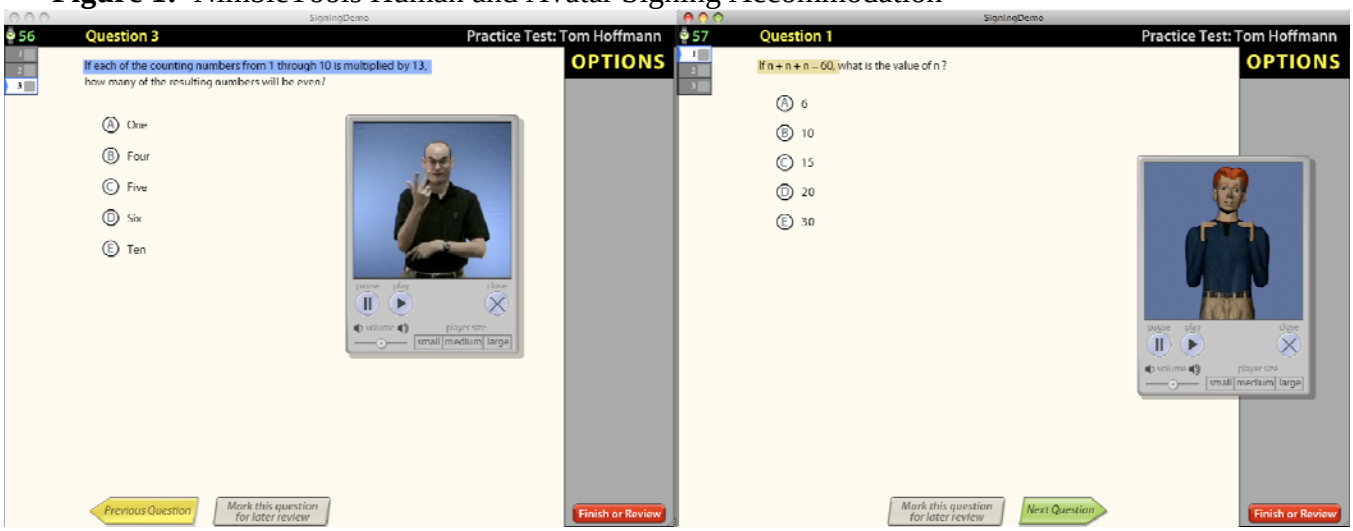
the human version was produced first. When producing the human version, a four-step process was employed. First, a script was created that documented how the item was to be presented in ASL. The script was reviewed and modified by experts who communicated in ASL and were familiar with the test content. Second, a human signer was recorded as he presented each item in ASL. Third, the recordings were reviewed by experts and suggested modifications were made for each item. Finally, the human was rerecorded based on suggestions provided by the experts.

Once the human recording was complete, the avatar was programmed to present the item in the same manner as the recorded human. To the extent possible, the two versions were as similar as possible. To assure that the highest quality signing was provided by the avatar, VCom3d, the company that developed the signing avatar software, was contracted to develop the avatar video footage. Nonetheless, because the avatar is not as expressive as an actual human, there were some differences in facial expressions and range of motion, with the human being more expressive and having a fuller range of motion.

Signing Accommodation and NimbleTools

As shown in Figure 1, the NimbleTools test environment presents test items one at a time. When the signing accommodation is activated, a window containing a video recording of signed presentation of the test item is displayed. Students may reposition the video window and may also adjust the size of the window. In addition, the video can be set to play automatically when a test item is presented on the screen. Alternatively, students can set the video to only play when they click on the play button or when they click on a specific block of text in the test question. Clicking on a block of text then plays only that portion of the video associated with the text. For students with partial hearing, students use a sound controller to adjust the volume of sound files that are synched with the signed video. Students may also use a "scrub-bar" that on the video window to move quickly or slowly forward or backwards through the video. In the full version of NimbleTools, students may also select to view the signed accommodation in either ASL or Signed English. For this study, however, students were limited to ASL presentation in order to standardize conditions across students. In addition, NimbleTools was set to play the video automatically when an item was displayed. Students could stop the video at any time and could click on blocks of text to view the signed interpretation at any time. Finally, to assure that students focused on the signed presentation, the sound files were not activated and the scrub bar was not displayed. These settings were made to increase the comparability of the delivery of signing across all students.

Figure 1: NimbleTools Human and Avatar Signing Accommodation



For this study, students were shown a short tutorial that explained how to use NimbleTools to perform a test and how to use the signing accommodation tool to view signed presentation of test items. Students were then presented with three practice test items that allowed them to become accustomed to using NimbleTools and the signing tool. Once they completed the practice items, students were presented with the 20-item test. Depending on their group, students either viewed the first 10 items using the recorded human or the avatar, and the second 10 items using the alternate signing method.

Once students were finished with the test, they were asked to complete a paper-based survey. The survey contained 15 items (shown in Table 2) and was designed to collect information about students experience using the signing tool, their attitudes about the use of the avatar as compared to a recorded human, and their preferences regarding the method used to provide a signing accommodation during testing. Each item presented students with a stem and four answer options that ranged from strongly agree to strongly disagree. We recognize that some survey developers support the use of a response scale that includes a mid-point because some respondents prefer not to select negative statements, yet do not necessarily agree with a statement. The mid-point allows them to indicate a non-positive response. However, we opted to employ a four-point scale without a neutral mid-point for three reasons. First, the provision of a neutral mid-point can result in unwanted equivocation (DeVellis, 1991). Second, a neutral mid-point may be used by respondents who truly do not have a preference and by those who prefer not to state their preference. Employing a scale with a non-neutral option removes this confusion. Third, and most importantly, given that state testing programs generally have limited funds to provide accommodations, we wanted to force students to state a non-neutral preference in order to assist testing programs in identifying a method that will be acceptable for the largest percentage of students.

Since many students' reading skills were below grade level, students were given the option of either reading the items themselves or having the survey items presented to them in ASL by the test proctor. While the option selected by students was not recorded, a substantial percentage of students opted to have the survey items presented to them in ASL. All students, recorded their answers on a paper-based survey. After students completed the survey items,

students were asked in a small group setting for suggestions on how the test delivery interface or the signing accommodation could be improved in the future.

Results

This study compares the effect that the use of a recorded human versus an avatar to provide a signing accommodation has on student preferences and performance on grade 8 mathematics test items. The first set of analyses presented below focus on students' responses to the survey items, which asked students about their use of the two modes of the avatar and the recorded human. The second set of analyses compares performance when items are presented by the recorded human or by the avatar. Finally, analyses are performed to examine the relationship between students' survey responses and their test performance. Specifically, we examine whether differences in students' comfort using NimbleTools were associated with differences in performance.

Student Survey

Table 2 displays the frequency of students selecting each response option for the 15 survey items. The majority of students reported that it was easy to perform the test on a computer (77.9%), that the tutorial was easy to use (84.9%), that the signing tools were easy to use (81.6%), and that they liked taking the test on computer (79.3%). A substantial percentage of students, however, reported that after completing the initial tutorial, they still were not sure how to take a test using a computer (36.7%). It is important to note, however, that the tutorial focused only on how to answer items on the computer, how to move between items, and how to end testing, but did not focus specifically on how to use the controls associated with the signing tools so that researchers could examine the extent to which the use of the tool was intuitive.

While the majority of students reported that it was easy to understand information presented by the avatar (59.7%), a higher percentage of students reported that the signing human was easy to understand (78.7%). Approximately half of the students reported that the avatar and human were equally effective for communicating test questions (53.5%). In addition, approximately one-third of students reported that they preferred the signing avatar rather than the signing human (43.4%). Interestingly, a small majority of students reported that it was distracting to have text in the test item highlighted while it was being communicated in ASL (56.3%).

When asked to compare the use of NimbleTools signing tools to other approaches to providing test accommodations, 59.7% reported preferring NimbleTools over watching a DVD containing signed items and 52.8% preferred NimbleTools over a live human.

Finally, students had mixed opinions about using a computer for a signing accommodation in the future. The vast majority indicated that they wanted to use a computer to take tests in the future (77.0%). A majority of students also indicated that they wanted to use the human signing video tools in the future (75.8%). A smaller percentage of students, however, indicated that they wanted to use the avatar video tools in the future (41.4%). Chi-square analyses indicated that none of the student responses to the survey items differed significantly between the two groups of students.

Table 2: Summary of Survey Responses

Item	Question	Strongly Agree	Agree	Disagree	Strongly Disagree
1	It was easy to perform the test on a computer.	25.6	52.3	17.4	4.7
2	The tutorial was easy to use	25.6	59.3	15.1	0.0
3	After using the tutorial, I still wasn't sure how to take a test on a computer	12.6	24.1	34.5	28.7
4	The signing tools were easy to use	35.6	46.0	13.8	4.6
5	It was easy to understand information presented by the signing avatar	21.8	37.9	28.7	11.5
6	It was easy to understand information presented by the signing human	54.0	28.7	13.8	3.4
7	The signing human and signing avatar were equally effective for communicating the test questions	18.6	34.9	32.6	14.0
8	I prefer the signing avatar tool rather than the signing human	10.3	24.1	40.2	25.3
9	While watching signed video, highlighting text in the question was distracting	21.8	34.5	27.6	16.1
10	I prefer using NimbleTools rather than watching a DVD of a person signing on a television or a DVD player	20.7	39.0	28.7	11.5
11	I prefer using NimbleTools rather than a live human signer to take a test	26.4	26.4	37.9	9.2
12	I liked taking the test on computer	43.7	35.6	14.9	5.7
13	I want to use a computer to take tests in the future	42.5	34.5	17.2	5.7
14	I want to use the signing avatar video tools to take tests in the future	18.4	23.0	40.2	18.4
15	I want to use the signing human video tools to take tests in the future	42.5	33.3	14.9	9.2

Test Performance

To examine whether the use of a recorded human versus a signing avatar affected test performance, we first compared the amount of time required for students to complete each test form. Students assigned to group 1 completed Form 1 using the recorded human and completed Form 2 using the signing avatar, while students assigned to group 2 received the two signing accommodations in the reverse order. As seen in Table 3, students assigned to Group 1 completed both forms faster than students assigned to Group 2. However, the differences in the time required to complete the test items was not statistically significant for either form.

Table 3: Comparison of Time to Complete Test for each Group

	Form 1		Form 2	
	Time (minutes)	**St. Dev.**	**Time (minutes)**	**St. Dev.**
Group 1 – Human-Avatar (n=49)	9.89	4.14	8.40	4.54
Group 2 – Avatar-Human (n=47)	10.23	5.81	9.32	4.76
Difference	0.33		0.91	
t-statics	0.33		0.96	
P	.75		.34	

Table 4 displays the mean and standard deviation for each test item and for the two item sets. Clearly, the items were generally difficult for the sample of students that participated in this study. Across all students, item difficulty ranged from .17 to .62. For the two item sets, students answered approximately 32% of the items correctly and the mean scores were only seven points higher than the chance level (.25). While there were some differences in performance between groups, there is no clear pattern to these differences. Moreover, the difference in mean scores expressed as an effect size were all small and ranged from .02 to .28, with the vast majority less than .15.

Table 4: Mean and Standard Deviation for Mathematics Test Items

	All Students (n=96)		Group 1 (n=49)		Group 2 (n=47)			
Item	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Difference	Effect Size
1	.62	.49	.57	.50	.68	.47	-.11	-.22
2	.35	.48	.35	.48	.36	.49	-.01	-.04
3	.22	.42	.22	.42	.21	.41	.01	.02
4	.24	.43	.27	.45	.21	.41	.06	.12
5	.34	.48	.33	.47	.36	.49	-.03	-.08
6	.28	.45	.27	.45	.30	.46	-.03	-.07
7	.20	.40	.14	.35	.26	.44	-.12	-.28
8	.40	.49	.43	.50	.36	.49	.07	.14
9	.29	.46	.27	.45	.32	.47	-.05	-.11
10	.20	.40	.22	.42	.17	.38	.05	.13
11	.53	.50	.47	.50	.60	.50	-.13	-.26
12	.39	.49	.41	.50	.36	.49	.05	.10
13	.42	.50	.43	.50	.40	.50	.03	.04
14	.17	.38	.18	.39	.15	.36	.03	.11
15	.26	.44	.24	.43	.28	.45	-.04	-.07
16	.25	.44	.20	.41	.30	.46	-.10	-.20
17	.34	.48	.33	.47	.36	.49	-.03	-.08
18	.29	.46	.35	.48	.23	.43	.12	.24
19	.35	.48	.35	.48	.36	.49	-.01	-.04
20	.18	.38	.18	.39	.17	.38	.01	.03
Form 1	3.15	2.14	3.06	2.09	3.23	2.22	-.17	-.08
Form 2	3.18	2.06	3.14	2.05	3.21	2.09	-.07	-.03

Independent samples t-tests were conducted to examine whether any differences in mean performance for individual items and for the two forms were statistically significant. As shown in Table 5, none of the differences were statistically significant.

Table 5: Results of Independent Samples t-Test for Test Items and Item Set Mean Scores

Item	Difference	Std. Error	*t*-static	*p*
1	-.11	.10	1.10	.27
2	-.02	.10	0.15	.88
3	.01	.09	0.14	.89
4	.05	.09	.60	.55
5	-.04	.10	0.36	.72
6	-.03	.09	0.35	.73
7	-.11	.08	1.38	.17
8	.07	.10	.66	.51
9	-.05	.09	0.58	.57
10	.05	.08	.66	.51
11	-.13	.10	1.24	.22
12	.05	.10	.46	.64
13	.02	.10	.24	.81
14	.04	.08	.45	.65
15	-.03	.09	0.35	.73
16	-.09	.09	1.06	.29
17	-.04	.10	0.36	.72
18	.11	.09	1.21	.23
19	-.02	.10	0.15	.88
20	.01	.08	0.17	.87
Form 1	-.17	.44	0.39	.70
Form 2	-.07	.42	0.17	.87

Test Score Performance of Higher Performing Students

Since many students performed at or below the chance level, a separate analysis of students who performed above 30% was conducted to examine differences in test performance when the signing accommodation was provided by a recorded human or an avatar. Of the 96 students who participated in the study, 45 answered 30% or more of the items correctly. As shown in Table 6, the mean scores for Form 1 and Form 2 were similar across groups and the differences were not statistically significant.

Table 6: Comparison of Test Scores for "Higher" Performing Students

	Form 1		Form 2	
	Mean	**St. Dev.**	**Mean**	**St. Dev.**
Group 1 (n = 24)	4.50	2.02	4.54	1.67
Group 2 (n = 21)	5.00	2.00	4.57	2.36
Difference	-.050		0.03	
t-statistics	0.83		0.05	
p	.41		.96	

Correlations Between Survey Responses and Test Performance

To examine the extent to which test performance correlated with students' opinions about the signing tools, the correlations between test scores and survey items were examined. As

shown in Table 7, after adjusting for multiple comparisons, there was a statistically significant correlation between test performance on both forms and student responses to only one survey item. This item (Q3) focused on the extent to which students were sure how to take a test on computer after completing the tutorial. As one might expect, students who reported feeling less sure how to perform a test after the tutorial tended to perform lower than students who reported feeling sure.

Table 7: Correlations Between Survey Item Responses and Test Performance

		Form 1	**Form 2**
1	It was easy to perform the test on a computer.	-.26	-.20
2	The tutorial was easy to use	-.19	-.12
3	After using the tutorial, I still wasn't sure how to take a test on a computer	.41*	.47*
4	The signing tools were easy to use	-.11	-.08
5	It was easy to understand information presented by the signing avatar	-.05	.08
6	It was easy to understand information presented by the signing human	-.29	-.23
7	The signing human and signing avatar were equally effective for communicating the test questions	-.14	-.13
8	I prefer the signing avatar tool rather than the signing human	-.12	-.08
9	While watching signed video, highlighting text in the question was distracting	.02	.10
10	I prefer using NimbleTools rather than watching a DVD of a person signing on a television or a DVD player	-.08	-.23
11	I prefer using NimbleTools rather than a live human signer to take a test	-.03	-.15
12	I liked taking the test on computer	-.19	-.12
13	I want to use a computer to take tests in the future	-.12	-.03
14	I want to use the signing avatar video tools to take tests in the future	-.12	-.03
15	I want to use the signing human video tools to take tests in the future	-.23	-.11

* significant at p = .05.

To remove the possible effect that non-familiarity with the test environment may have on students' performance, a separate analysis was conducted that focused only on those students who reported feeling comfortable taking a test on computer after completing the tutorial. As seen in Table 8, both groups performed nearly identical on Form 1. For Form 2, Group 2, which used

the recorded human, performed better than Group 1, which used the signing avatar. This difference, however, was not statistically significant.

Table 8: Comparison of Test Scores for Students Comfortable Taking Test on Computer

	Form 1		**Form 2**	
	Mean	**St. Dev.**	**Mean**	**St. Dev.**
Group 1 (n = 30)	3.70	2.37	3.77	2.08
Group 2 (n = 25)	3.72	2.73	4.12	2.37
Difference	-.02		-.35	
t-statistic	0.03		0.59	
p	.98		.56	

Discussion

The equitable provision of high-quality signing accommodations for students who are deaf or hard-of-hearing presents a major challenge for large-scale testing programs. Past research documents many problems that schools and testing programs encounter when providing signing accommodations, including the use of interpreters who are unfamiliar with test content, the inability to familiarize interpreters with test content prior to test administration due to security concerns, provision of signing accommodations in a group rather than individualized setting, and differences in the way that signers interpret and present test content.

To overcome some of these shortcomings, some testing programs deliver the signing accommodation on video-tapes or DVD. While this strategy standardizes the provision of the accommodation across students, it also has shortcomings. Among them are the need to use remote controls to fast-forward and rewind through video footage, the delivery of the accommodation in a group, rather than an individual setting, and a physical separation between the test booklet and the presentation of signed interpretations of the test content.

A recent movement towards computer-based test delivery holds promise to standardize the provision of signing accommodations. Computer-based delivery of signing accommodations also places the signed presentation in close proximity to the test content, simplifies the process of replaying sections of video, and allows students individualized control of the signed presentation. In addition, computer-based delivery also holds promise to decrease the cost of creating signed video by using avatars rather than recordings of a human interpreter. The use of avatars also holds potential to allow students to further tailor the provision of signing accommodations to meet their specific needs by activating or deactivating lipping, audio, and by switching between ASL and Signed English. However, despite these potential advantages, concerns have been raised that the use of avatars may adversely affect the quality of signed accommodations, which in turn may have a negative effect on student test performance.

The study presented here employed a randomized trial to compare the effect that a computer-based provision of the signed accommodation using a recorded human versus a signing avatar had on students' attitudes about performing a mathematics test and on their actual test performance. This study found that students generally reported that it was easy to perform a mathematics test on computer, and that both the recorded human and the signing avatar tools were easy to use and to understand. Students also reported a strong preference for performing future tests on computer, and generally preferred using the recorded human and the avatar for future tests rather than a DVD. While students less than half of the students reported that they preferred the signing avatar rather than the recorded human presented on computer, this

preference was unrelated to test performance. The use of the recorded human and the avatar did not have affects on either the amount of time required to complete the test items or on students' performance on the test items.

During focus groups that occurred after students completed the test and the survey, students were very positive about their experience using the NimbleTools signing tools. Several students had performed the Massachusetts state test the prior spring and stated that they strongly preferred the NimbleTools signing tools over the DVD version that was used for the actual state test. And, although they generally preferred the human over the avatar, this sub-group of students all stated they would rather use the NimbleTools avatar rather than the DVD, primarily because it gave them complete control over playing the signed video.

Students also noted a few aspects of the signing tool that could be improved. First, many students stated that the highlighting of text on the screen as it was being presented in ASL was distracting. While they liked that they could click directly on a block of text in order to play the portion of the video associated with that text, they found that highlighting of blocks of text while the video was playing such that the highlighting moved with the signed video caused their eyes to move away from the video to focus on the text. Second, several students noted that automatically playing the signed video when a new item is displayed on the screen was distracting. These students would prefer to read the item first and then play the video when needed. Although the NimbleTools signing tool includes a feature that allows students to control whether or not video is played automatically when it is loaded, this feature was deactivated for this study to assure that all students were exposed to the video. Clearly, this feature is important to include for an actual test administration and for future studies. Finally, some students noted that the avatar's movements were not as lifelike as the humans. While most students felt that this was not problematic, a few students stated that their unfamiliarity with the avatar and its more "jerky" movements were distracting. It will be important to monitor whether this issue abates as the quality of avatars continues to improve. In addition, it is important to explore whether familiarizing students with the avatar prior to testing helps acclimate students to the avatar's movements and thus help them feel more comfortable viewing signed interpretations presented by the avatar.

Finally, for this study, the tutorial provided an introduction to taking a test on computer using text-based explanations. As noted above, several students who participated in this study had reading skills that were below grade level. Several of these students complained that the tutorial was difficult to understand and indicated a preference to have the tutorial presented in ASL. In addition, one teacher who assisted with the study suggested that future tutorials include a signing option. This shortcoming may partially explain why some students did not feel comfortable taking a test on computer after completing the tutorial. It also highlights the importance of providing the signing accommodation for all aspects of a test for future studies, including directions and post-test surveys.

Despite these observations and comments by students, the reaction to using a computer-based signing tool was positive. Furthermore, the use of the human and avatar appeared to have the same effect on student performance. Nonetheless, this study had several shortcomings. First, a convenience sample that included students that ranged from eighth to eleventh grade was used. Future studies should focus more narrowly on students within a single grade level or grade levels in close proximity. Second, the signing accommodation is intended to assist students who have difficulty reading. Given this, it would be useful to collect information about students' reading skills and examine whether the use of a human or an avatar differs for students at different levels

of reading proficiency. Third, this study focused specifically on a recorded human and an avatar, but did not provide any insight into whether either method was more effective than either a live human or a DVD version of a recorded human. Similarly, the study did not provide any insight into whether the signing accommodation resulted in scores that differed from those that students would obtain if no accommodation was provided. Future studies should be conducted to compare the use of computer-based recorded human, avatars, DVD versions of a recorded human, live humans, and no accommodations. Future studies should also examine the degree to which the use of live humans affects the consistency with which the signing accommodation is provided across students. Finally, a cost analysis should be performed to develop a solid understanding of the costs associated with using live humans, recorded humans, and avatars. With this full set of information, sound decisions about the effects on the equitable provision of high-quality signing accommodations, and the associated costs, can be made.

Despite these shortcomings and the need for further research, the study presented here provides preliminary evidence that students prefer a signing accommodation that is integrated into a computer-based test delivery system over watching a recording of a person displayed on a television of DVD player. In addition, this study provides preliminary evidence that, despite students' preference for a recorded human over an avatar, the use of an avatar does not have an adverse effect on test performance. Collectively, these findings hold promise for the future of computer-based testing to provide signing accommodations in an equitable, high-quality, yet individualized manner that meets the needs of students who are deaf or hard-of-hearing.

Acknowledgements

The authors wish to thank teachers at the two schools whose students participated in this study. This material is based upon work supported by the Institute of Education Sciences under Grant No. ED-08-CO-0045. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the Institute of Education Sciences.

References

Abedi, J., Hofstetter, C., Baker, E., & Lord, C. (2001). *NAEP math performance and test accommodations: interactions with student language background. CSE technical report* No. CSETR536). U.S.; California: University of California.

Bielinski, J., Thurlow, M., Ysseldyke, J., & Fieidebach, M. (2001). *Read-aloud accommodations: Effects on multiple-choice reading and math items.* Minneapolis, MN: National Center on Educational Outcomes.

Cawthon, S. & the Online Research Lab (2006). National Survey of Accommodations and Alternate Assessments for Students who are Deaf or Hard of Hearing. *Journal of Deaf Studies and Deaf Education*, 11(3), 337-359.

Clapper, A. T., Morse, A. B., Thompson, S. J., Thurlow, M. L. (2005a). *Access assistants for state assessments: A study of state guidelines for scribes, readers, and sign language interpreters* (Synthesis Report 58). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved [today's date], from the World Wide Web: http://education.umn.edu/NCEO/OnlinePubs/Synthesis58.html.

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

Dalton, B., Kennedy, M., Lutz, M. & Schleper. (2003). Shared Reading Project: Chapter by Chapter, a presentation to the Board of Trustees at Gallaudet University.

DeVellis, R. F. (1991). *Scale development: theory and applications*. Newbury Park, CA: SAGE Publications, Inc.

Driscoll, D. P. (2007). *Requirements for the participation of students with disabilities in MCAS*. Malden, MA: Massachusetts Department of Education.

Elliot S.N., Kratochwill, T.R. & Schulte, A.G. (1999). *Assessment accommodations checklist.* Monterey, CA: CTB/McGraw Hill.

Horvath, Leah S., Kampfer-Bohach, S. & Farmer-Kearns, J. (2005). The use of accommodations among students with deafblindness in large-scale assessment systems. *Journal of Disability Policy Studies*, 16(3), 177-187.

Individuals with Disabilities Education Act, Pub. L. No. 108-446, 118 Stat. 2647 (2004).

Johnson, E., Kimball, K., & Brown, S.O. (2001). American sign language as an accommodation during standards-based assessments. *Assessment For Effective Intervention, 26*(2), 39-47.

Lazarus, S. S., Thurlow, M. L., Lail, K. E., Eisenbraun, K. D., & Kato, K. (2006). *2005 state policies on assessment participation and accommodations for students with disabilities* (Synthesis Report 64). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved December 8, 2008, from: http://education.umn.edu/NCEO/OnlinePubs/Synthesis64/.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).

Russell, M., Higgins, J., & Hoffmann, T. (2009). Meeting the Needs of All Students: A Universal Design Approach to Computer-Based Testing. Innovate, 5(4). Retrieved May 15, 2008 from http://www.innovateonline.info/index.php?view=article&id=676&action=synopsis.

Sireci, S.G., Li, S., & Scarpati, S. (2003). *The effects of test accommodations on test performance: A review of the literature* (Research Report 485). Amherst, MA: Center for Educational Assessment.

Thompson, S. & Thurlow, M. (2001) *State Special Education Outcomes: A report on state activities at the beginning of a new decade.* http://education.umn.edu/nceo/onlinepubs/2001statereports html.

Tindal, G., Heath, B., Hollenbeck, K., Almond, P., & Harniss, M. (1998). Accommodating students with disabilities on large-scale tests: an empirical study of student response and test administration demands. *Exceptional children, 64*(4), 439-450.