MARC MARSCHARK, PATRICIA SAPERE, CAROL CONVERTINO, ROSEMARIE SEEWAGEN, AND HEATHER MALTZEN

Comprehension of Sign Language Interpreting: Deciphering a Complex Task Situation

FOR MANY DEAF and hard of hearing people, sign language interpreting is essential to their participation in various educational programs; daily activities relating to employment, public services, and interactions with local businesses; and medical and mental health services (e.g., Harmer 1999; Leigh and Pollard 2003; Winston 1994). Unfortunately, there are not enough interpreters available to meet the demand. Baily and Straub (1992) estimate that to satisfy current needs, the United States alone would need approximately 25,000 interpreters. In 1995, however, the Registry of Interpreters for the Deaf (RID) reported only about half that number (RID 1995), and current information from RID (www.rid.org, accessed February 12, 2003, and RID, personal communication, February 12, 2003) reveals that among their 20,000 members, there are fewer than 3,500 certified interpreters and 2,600 associate members. Although these statistics are from only a single country, the United States appears far ahead of

Marc Marschark, Patricia Sapere, Carol Convertino, and Rosemarie Seewagen are employed at the National Technical Institute for the Deaf, a college of Rochester Institute of Technology. Marc Marschark also has an appointment at the University of Aberdeen. Heather Maltzen is employed at Rochester School for the Deaf.

most others in terms of the quality and availability of interpreting services.

Beyond the shortage of interpreters, frequent mismatches occur between interpreter skills and the needs or preferences of deaf individuals, and interpreters are sometimes assigned to situations for which they are unprepared or unqualified (see Schein, Stewart, and Cartwright 1998; Seal 1998). Indeed, a 1987–1988 study by the National Association of the Deaf (NAD) found that 70 percent of the deaf individuals participating were dissatisfied with interpreting quality. Such situations leave both interpreters and their clients dissatisfied and lead interpreter educators to look for new methods to improve the flexibility and skill of new members of the profession (Monikowski and Winston 2003; Roy 2000; Seal 1998).

Despite such challenges to interpreters and the deaf and hearing clients they serve, research devoted to the outcomes of interpreting is remarkably scarce. Numerous studies have documented the need for sign language interpreters, processes thought to be involved in effective interpreting, and alternative means of evaluating interpreting skills (see Monikowski and Winston 2003). However, little information is available about how variables that interpreters and clients believe influence interpreting (e.g., interpreter background knowledge, interpreter familiarity with clients and with to-be-interpreted material) actually influence comprehension of interpreted material. In the case of educational interpreting, there is almost a complete lack of knowledge concerning how various interpreting variables might interact with characteristics of deaf students or with various learning situations such as different ages, grade levels, or class content (Kluwin and Stewart 2001; Stewart and Kluwin 1996). Medical and legal interpreting are acknowledged as requiring special skills, but the demands of university versus K–12 interpreting and science, technology, engineering, and mathematics (STEM) versus social science and liberal arts interpreting are still in need of investigation (Harrington 1999; Lang 2002; Napier 2002; Redden, Davis, and Brown 1978).

One primary concern for interpreters and interpreter educators, as well as for deaf people, is how to best match the style or mode of interpreting to the clients' preferences and signing skills (e.g., ASL or various forms of English-based sign communication).[1] This issue is

particularly important in the classroom, where students often have heterogeneous language backgrounds and diverse signing skills (Harrington 1999; Napier 2002; Napier and Barker 2003). In such settings, the educational interpreter faces the challenge of trying to sign in a way that is congruent with the students' skills or, more commonly, in a manner satisfactory for the wide range of skills represented in the classroom. Among the few studies that have considered this issue, however, there is considerable disagreement.

Fleischer (1975), for example, investigated the comprehension of a lecture by deaf high school students under four conditions: interpreting (ASL) after interpreters had been supplied with background information about the lecture, interpreting without background information, transliteration with background information, and transliteration without background information. No main effect of presentation condition was obtained, but Fleischer found that, overall, the interpreted conditions led to significantly better comprehension than the transliteration conditions. Correlations examining what little demographic information Fleischer had for his participants suggest that students' language fluencies might interact with the mode of communication, but he did not have sufficient data to evaluate that possibility fully.

In a related study, Livingston, Singer, and Abramson (1994) found that, when deaf community college students were assigned to transliteration or ASL interpretation conditions, those in the latter group showed higher overall comprehension scores. Prior to testing, students were "interviewed individually by either two or three bilingual Deaf adults to determine their Sign Preference and Communicative Competence in expressing and receiving either ASL or English-like signing" (178). Among students who had seen a lecture interpreted (in ASL), those who were designated as ASL-oriented showed a significant advantage relative to students who were designated as oriented toward English-like signing. A reliable advantage was not obtained for transliteration of the lecture by the students who were designated as oriented toward English-based signing. Further, when a narrative presentation rather than a lecture was interpreted, neither comparison was reliable, making it difficult to draw any firm conclusions. The study by Livingston et al. also appears to be confounded

by the fact that different interpreters were involved in each of fifteen testing sessions. There were also large comprehension differences across those sessions, ranging from 50 to 84 percent in the six ASL narrative groups, the only condition for which such information is provided. It is impossible to know what effects, if any, such differences might have had on their results, but such variables clearly need to be carefully controlled if experimental results are to be considered valid and reliable. Just as importantly, with an overall score of approximately 62 percent, the students in that study showed rather poor comprehension.

Murphy and Fleischer (1977) replicated Fleischer's (1975) study, comparing interpreting and transliteration with groups of deaf students who reported preferring one mode or the other. In contrast to the suggestive evidence offered by Fleischer and Livingston et al., Murphy and Fleischer found no differences in comprehension due either to mode of communication or to communication preference (or their interaction). It is noteworthy, however, that Murphy and Fleischer used a written, multiple-choice comprehension test. Although they did not examine the possibility of a confound with reading levels, the documented reading challenges that deaf students experience (e.g., Traxler 2000) suggest that the issue is worthy of consideration. In the Murphy and Fleischer study, for example, it is unclear whether participants failed to understand the interpreter in one or both conditions or simply failed to understand the written test questions (Hatfield, Caccamise, and Siple 1978).

If one accepts their findings at face value, the studies by Fleischer (1975) and Livingston et al. (1994) might be seen as supporting the use of a natural sign language (ASL) over the combination of signs with English grammar in transliteration or simultaneous communication (in direct instruction), an argument mounted in theoretical terms by Johnson, Lidell, and Erting (1989) and others. Importantly, however, the studies that proponents of that position usually cite have generally involved parents who were beginning signers or teachers who had only minimal signing skills and were documented within the studies to be poor at communicating spoken information in sign (cf. Leigh 1995). There is some contrasting evidence that transliteration can lead to high levels of comprehension and learning among deaf students when done by a skilled transliterator (e.g., Caccamise

and Blasdell 1977; Napier 2002; Newell 1978), but the issue is clearly in need of resolution.

With regard to student preferences rather than comprehension per se, Napier and Barker (2004) recently conducted a qualitative study in which they examined deaf university students' preferences for "free interpretation" (primarily Auslan) versus "literal interpretation" (primarily transliteration). Their results clearly indicate that even students who preferred Auslan in social and general academic settings wanted transliteration in more technical courses so that they could acquire the same vocabulary as hearing peers. That consensus was consistent with Napier's (2002) finding that university-level interpreters routinely codeswitch in order to provide deaf students with information necessary for their academic success.

The preceding studies focus on the relative merits of different modes of presentation on comprehension by different groups of deaf students. However, it is important to emphasize that the underlying issue is a more basic one. By means of a written comprehension test, Jacobs (1977) found that in a classroom setting, hearing students uniformly showed significantly greater comprehension of content relative to deaf students who relied on interpreting. While the possible confounding of comprehension with (test) reading ability was not considered, a clear possibility remains that deaf students generally comprehend less than hearing peers in mainstream classrooms (even when language barriers appear to be removed), due to inefficient learning strategies or earlier educational and language experiences (e.g., Lang 2002; Marschark, Lang, and Albertini 2002; Richardson, MacLeod-Gallinger, McKee, and Long 1999). It is customarily assumed that providing deaf students with access to lectures and classroom discussion via interpreting or transliteration offers them learning opportunities comparable to those of hearing students, but there is little information available to support that argument.

The present study was therefore designed to begin a carefully controlled inquiry into the factors that contribute to classroom learning through sign language interpreting. Three experiments addressed the basic questions of the extent to which deaf students understand classroom communication presented in sign language, how comprehension is related to their content knowledge as well as language skills and preferences, and the possible role of reading ability and

written comprehension tests in assessing deaf students' learning of classroom material.

Experiment 1

Experiment 1 examined the sign language skills and preferences of deaf students and the mode of production (interpreting versus transliteration), as in the studies of Fleischer (1975) and Livingston et al. (1994). Extensive demographic information concerning the student participants was evaluated in order to provide more clarity and confidence in the results.

Given the contradictory findings from previous research, predictions for this experiment are not as straightforward as one might expect. Many deaf students and some investigators in the field would expect that natural sign languages offer the most effective means of communication with deaf students and would thus predict that interpretation would lead to better overall comprehension than transliteration (e.g., Johnson, Liddell, and Erting 1989; Lang, McKee, and Conner 1993). Given the English orientation of education in the United States (through reading, even if not through the mode of communication in the classroom) and the fact that few sign language interpreters are truly fluent in ASL, it appears equally likely that transliteration could prove as effective as interpreting (or even more so) in fostering comprehension (see Mayer and Akamatsu [1999] for a discussion related to literacy). More likely than either result, however, would be an interaction whereby students with greater skills in one mode or another would perform better when those skills match the mode of interpreting than when they mismatch, as Livingston et al. (1994) suggest.

Method

Participants. The participants were thirty-eight deaf students attending the Rochester Institute of Technology (RIT), enrolled either in the National Technical Institute for the Deaf (NTID) or one of the other colleges of RIT, and ten hearing RIT students. They were recruited via personal contact and paid for their participation. Demographic data were available for most of the deaf students. (Where databases were not complete, the number of cases for which data

were available is indicated in the parentheses below.) Hearing thresholds (35) ranged from 55 dB to 120 dB in the better ear, with a mean threshold of 99 dB (SD = 16 dB). Sixty-seven percent (30) of the deaf students reported using hearing aids, and one reported having a cochlear implant, although this study did not evaluate their use. In addition to information on language skills collected in the context of this study, data (30) were available from the Language and Communication Background Questionnaire (LCBQ), which most deaf students at RIT fill out when they first enroll. NTID now employs the LCBQ rather than face-to-face communication interviews to obtain information on student sign language and spoken language skills because it is more efficient and has been found to correlate approximately .80 with interview assessments.

Materials. For the purposes of this study, a nonteaching staff member at RIT was recruited to make a short presentation about her hobby of collecting antique toys. The five-minute presentation was given in spoken language only and was videotaped in a television studio. The presenter spoke at a rate of approximately 175 words per minute (wpm), well within the range of interpreted lectures presented in university classrooms. The presenter was blind with respect to the goals of the study or the intended use of the videotape until the taping was completed.

The tape was watched several times by a professional interpreter with almost twenty years of professional interpreting experience in educational, vocational rehabilitation, and other settings (ten years in the NTID/RIT context); RID certification; and exposure to sign language at home, having been raised with a deaf sibling. The interpreter practiced interpreting and transliterating the presentation and was eventually videotaped while signing one version of each. To increase accuracy, both versions were taped a second time, and the second versions were those that were actually used.

For the purposes of testing comprehension, the investigators watched the videotape and created ten multiple-choice questions, all with four plausible answer options. The questions ranged from central information (e.g., "What does the woman collect?") to details relevant to her presentation (e.g., "Where does she keep the items

she collects?''). Subsequent examination of the videotape (after testing) by two skilled, certified interpreters confirmed that both the interpretation and transliteration of the presentation were comprehensible and appropriate, given the test. One question was thought to be biased, however, in that the ASL interpretation mentioned its correct answer only elliptically. That question was excluded from analyses described later for both deaf and hearing students (leaving nine questions).

Procedure. Testing was done in small groups, with students seated as they normally would be in a classroom with a video projection system. They were told simply that they would see a short, interpreted presentation about collecting (using a split screen) and then would be asked questions. Half of the deaf students (nineteen) saw the ASL version of the presentation, and half saw the English transliteration, according to the testing session in which they participated (alternating sessions until the last two, when the sizes of the two groups were balanced). Hearing students saw a version of the videotape without interpreting. To eliminate recency effects, a brief delay followed the presentation, during which the comprehension test was distributed and students filled in identifying information. The instructions were given using simultaneous communication to the deaf students and spoken language to the hearing students; instructions were also given in written form to all of the students. Participants then completed the comprehension test.

After the comprehension test, the deaf students answered four additional questions about their language preferences and skills, intended to complement the LCBQ questions. Those four questions involved marking answers on seven-point Likert scales: (1) Overall, I prefer to use: sign language [rating scale] spoken language; (2) Overall, I prefer to use: ASL [rating scale] signed English; (3) Overall, my ASL skills are: bad [rating scale] good; and (4) Overall, my signed English skills are: bad [rating scale] good. These questions provided both information that is slightly different from that obtained by the LCBQ and also a measure of language skills at the time of testing, whereas the LCBQ had been administered at the time of RIT entry. Neither the LCBQ nor the questions used here were intended as

definitive, precise assessments of student language skills, but, looking ahead, they provided estimates sufficient for the present purposes.

Testing in Experiments 1 and 2 was conducted by the first two authors. The first author has an advanced rating on the Signed Communication Proficiency Interview; the second author is the interpreter described earlier.

Results and Discussion

Unless noted otherwise, throughout this article all and only those results reported were significant at the .05 level. Comprehension scores are expressed as proportion correct.

Pairwise t-tests (with a Bonferroni adjustment and harmonic degrees of freedom) indicate that the interpretation and transliteration groups of deaf students did not differ reliably in their performance, $t(36) = .95$, while both groups scored significantly below the hearing students, $t(20.33) = 3.61$ and $t(21.25) = 4.39$. Mean scores for the three groups were as follows: interpretation .62 (SD = .19), transliteration .56 (SD = .20), and hearing students .87 (SD = .17). Five hearing students and one deaf student obtained perfect scores on the comprehension test.

Considering only the deaf students, those in the two conditions did not differ significantly on any of the communication or demographic variables, according to two-tailed t-tests. Multiple regression analyses for the entire group of thirty-eight deaf students and for the separate groups, using the communication and demographic variables, revealed no reliable predictors of comprehension. These results suggest that the two groups were comparable demographically and that the two interpretations are comparable and unbiased.

The results of this experiment indicate that regardless of students' preferences and skills with regard to signed communication, they understood interpreting and transliteration equally. However, their overall average comprehension score was only 59 percent, reliably lower than the 87 percent of the hearing students. This finding may reflect a floor effect in performance that did not allow us to discern true differences in performance between the interpreting and transliteration conditions as a function of student sign skills and preferences, although the fact that performance was well above chance (25

percent) would seem to rule out this possibility. Alternatively, the difference in performance could be the consequence of some global loss of information or cohesion in interpreted settings (i.e., mediated instruction versus direct instruction; Harrington 1999; Jacobs 1977; Lang 2002).

Given deaf students' documented challenges in reading (e.g., Traxler 2000), a more likely explanation for the relatively poor performance of deaf students in Experiment 1 is that although they might have understood the interpreting, they may have had more difficulty with the written evaluation of their comprehension (Hatfield, Caccamise, and Siple 1978). The use of a written evaluation is the most ecologically valid way of conducting such research because that is the way course material is normally tested. As noted earlier, however, this potential confound has also been present in previous studies and clouds the interpretation of these and previous results.

In order to evaluate the possibility of a confound with reading abilities in this experiment, scores were obtained from reading-related tests required for admission to RIT. Although RIT normally requires students to take the American College Test (ACT), not all of the deaf students participating in this experiment had done so; various students' records included scores on the ACT (29), the Scholastic Aptitude Test (SAT) (verbal portion) (5), the Michigan (36) and California (36) reading tests, and NTID's internal reading (33) and writing (34) tests, which are used for course placement. Analyses indicate that, by chance, the group receiving the ASL interpretation had a lower average reading score than the group receiving the English-based interpretation, both on the California test (7.76 SD = 1.06 and 8.77 SD = 1.56, respectively), $t(34) = 2.28$, and on the NTID reading test (91.67 SD = 17.54 and 113.44 SD = 27.95, respectively), $t(31 = 2.62)$. However, reading scores did not predict comprehension performance in correlational or regression analyses for either the whole sample or the two subsamples. Further, a one-way analysis of covariance, holding reading ability constant, does not change the results (reading scores were not available for hearing students).

The results of Experiment 1 therefore indicate that deaf students who vary considerably in their preferences for and skills in ASL and English-based signing nonetheless understand interpretation and

transliteration of simple lecture presentations in the two modes equally. The only reliable correlate of overall comprehension performance is the self-rated skill in English-based signing, which is associated with poorer performance, but that finding was not obtained in the group that received the transliteration. Although a link between a preference for transliteration and poorer performance would be consistent with suggestions that English-based signing generally does not offer an effective mode of instruction (Johnson, Liddell, and Erting 1989), the lack of any advantage for students who received the interpretation, regardless of their reported ASL skills, leaves some uncertainty with regard to firm conclusions from this experiment. Experiment 2 was therefore designed to offer a replication of this experiment while also exploring the possibility that while ecologically valid in mainstream classrooms, the use of written comprehension tests in research of this sort might provide an overly conservative view of deaf students' understanding of classroom interpretation.

Experiment 2

Experiment 2 replicated Experiment 1 in essentially all respects, except that signed tests were used instead of a written comprehension test. Although not as ecologically valid as Experiment 1, and perhaps not very practical given deaf students' varied sign language skills and the realities of mainstream classrooms, this methodology was intended to provide an optimal estimate of deaf students' comprehension performance in an interpreted setting.

Method

Participants. Forty deaf students participated. All were drawn from the same population as in Experiment 1 in the same way and were again paid for their participation. None had participated in the earlier experiment. They were randomly assigned to the ASL and English transliteration conditions, as in Experiment 1. Hearing students were not tested.

Design and Materials. The design and materials were the same as in Experiment 1. The only difference was that the person who had initially interpreted the presentation was videotaped interpreting the

multiple-choice questions used in Experiment 1 in the same television studio that had been used in Experiment 1.[2] Following practice, the first author read each question while the second author interpreted it in either ASL or transliteration, yielding both audio and video tracks for presentation (comparable to the original lecture).

Procedure. The procedure was the same as in Experiment 1, except that students who received the interpreted presentation then saw interpreted questions and those who received the transliteration saw transliterations of the questions. Each question was signed and spoken twice, including all four of the response alternatives. Students circled the letter *A, B, C,* or *D* on an answer sheet, corresponding to the letter of their selected answer. Inadvertently, the four supplementary communication questions were not asked during the testing session proper for about half of the participants. The questions were emailed (several times) to those students afterward, but questionnaire data were available for only thirty of the forty participants.

Results and Discussion

Overall, performance in the interpreting condition (.61 SD = .20) and the transliteration condition (.59 SD = .21) did not differ reliably by a two-tailed t-test. One student received a perfect score on the comprehension test. All of the analyses conducted in Experiment 1 were repeated for Experiment 2, and no reliable differences were obtained as a function of any of the variables examined.

Comparing Experiments 1 and 2, it is obvious that giving students a signed test after a presentation did not improve performance. The overall correct performance score was only .60 SD = .20. A two- (interpretation or transliteration) by-two (written or signed test) analysis of variance indicates that neither factor was significant, both Fs ($1,74 < 1.0$). Even though we are cautious about accepting null results, the replication of the findings from Experiment 1, together with similar findings in similar, if less well-controlled previous studies, suggests that the results of these experiments are both valid and reliable. Nevertheless, a third experiment was conducted to provide a more rigorous evaluation of the utility of interpreting and transliteration in the classroom.

Experiment 3

Two aspects of Experiments 1 and 2 potentially limit the generality of the findings. First, the presentation was on a topic that was undoubtedly unfamiliar (and perhaps uninteresting) to the students and was given by someone who was not an instructor. Although this was true for hearing as well as deaf students, greater heterogeneity in deaf students' world knowledge (e.g., McEvoy, Marschark, and Nelson 1999) may have biased the task against them. In Experiment 3 an RIT faculty member provided an introductory-level presentation more akin to a real classroom lecture, and students' prior content knowledge was assessed. Second, the previous two experiments included the same, skilled interpreter, who was also one of the investigators. Experiment 3 involved a different, highly skilled interpreter and also included several more methodological controls to provide greater confidence in the results.

Method

Participants. Forty-eight deaf students and 20 hearing students were recruited from the same population as in the previous experiments. Short questionnaires completed prior to testing provided information on students' sign language skills and preferences and were used to assign them to separate groups for testing purposes. Twenty-five of the deaf students indicated stronger skills and preferences for ASL, and 23 indicated stronger skills and preferences for English-based signing.

Materials. RIT faculty were recruited to provide short lectures in their content area for use in the larger project on accessing technical education through sign language interpreting. The lecture used in this experiment was an introductory-level lecture by a physics professor on "why we see our breath on cold mornings." No visual aids were used. The eight-minute lecture was presented and digitally recorded twice, and the second recording was used. The instructor also provided information for the development of a six-question pretest of related content knowledge and the twelve-question comprehension test, both of which were in multiple-choice format. The instructor's spoken language speed was approximately 180 wpm.

After practicing with the videotaped lecture, interpreted and transliterated versions of the lecture were produced by a highly skilled, certified interpreter with more than twenty years of experience; all technical terminology was fingerspelled in both versions (Napier and Barker 2003; Napier 2002). The interpreter uses ASL as the primary home language (deaf spouse) and is acknowledged to have superior interpreting and sign language skills, particularly in ASL. Two other highly skilled, certified interpreters carefully examined the interpretations in the context of the comprehension test. We then made minor adjustments to the test to ensure that the questions were entirely fair and balanced given the two interpretations.

A revised communication questionnaire, provided in the Appendix, was given at the end of the experiment. In addition, students were asked to list the number of physics courses that they had taken at the university level and to predict the number of questions they would get correct on the comprehension test (0–3, 4–6, 7–9, or 10–12 questions).

Procedure. Participants were tested in groups and paid for their participation. The experimenters arranged for meetings with student groups (e.g., fraternities, sororities), where they administered the short questionnaire concerning sign language skills and preferences. On the basis of students' responses, each group was divided into ASL-oriented and English-based sign-oriented groups, which were tested separately. Over the course of the several testing sessions, approximately half of the deaf students who identified themselves as having each orientation participated in the interpretation condition, and approximately half participated in the transliteration, yielding a two- (sign orientation: ASL or English-based sign) by-two (received: interpretation or transliteration) design with ten to fifteen students per cell.

In all other respects, testing was identical to that in Experiment 1 (e.g., a written comprehension test; hearing students saw a version of the lecture without interpreting).

Results and Discussion

An initial two-way analysis of variance involved the four groups of deaf students in the two-by-two design described earlier using pretest scores as the dependent variable. No effects of the sign orientation

factor, $F(1,44) = 3.16$, the interpretation vs. transliteration factor, $F(1,44) = .001$, or their interaction, $F(1,44) = .221$, $MSe = 1.09$, were obtained. Combining the deaf students into a single group and comparing their pretest performance to that of the hearing group, in contrast, yielded a reliable difference, $t(66) = 3.81$, as the hearing students obtained higher scores on the pretest of content knowledge (see table 1). The groups did not differ in the number of physics courses they had reported taking, and the number of physics courses taken was unrelated to pretest performance, $r(47) = .08$, suggesting that the lecture was of a sufficiently introductory level (i.e., was not dependent on how much physics they knew).

Four deaf students (8 percent) and five hearing students (25 percent) scored perfectly on the comprehension test. Comprehension scores were analyzed in two ways. First, a two-way analysis of variance of the two-by-two design involving deaf students was conducted, using comprehension scores as the dependent variable. No effects of the sign orientation factor, $F(1,44) = .03$, the interpretation vs. transliteration factor, $F(1,44) = .62$, or their interaction, $F(1,44) = .001$, $MSe = 2.76$, were obtained. Combining the deaf students into a single group and comparing their comprehension scores to those of the hearing group again yielded a reliable difference, $t(66) = 4.62$, as the hearing students scored higher on the comprehension test than their deaf peers (see table 1).

Given the apparent differences in prior content knowledge, a second analysis compared comprehension test scores using an analysis of covariance, in which pretest scores were the covariate. A one-way analysis with five groups (four deaf, one hearing) yielded a reliable main effect of group, $F(4,62) = 2.77$, $MSe = 2.27$, but pairwise post hoc comparisons (with a Bonferroni adjustment) did not yield any significant differences among the groups when prior knowledge was controlled.[3] Correlations between the number of questions correct and students' predictions in that regard indicate that, as a group, the deaf students were quite poor in predicting their performance, $rs(9–14) = .26–.41$, n.s., whereas the hearing students were able to predict their performance reliably, $r(19) = .57$.

Demographic variables from the deaf students' records were analyzed as in the previous experiments, yielding no significant differences among the groups. One could argue, however, that

Table 1. Mean Scores and Standard Deviations on Content Knowledge Pretest and Comprehension Test in Experiment 3

Group	*N*	Pretest	Comprehension Test	Controlling for Pretest
ASL skilled—ASL received	15	.37 (.08)	.74 (.12)	.75
ASL skilled—transliteration* received	10	.38 (.11)	.78 (.11)	.78
English signing skilled—ASL received	13	.43 (.09)	.75 (.16)	.75
English signing skilled—transliteration received	10	.42 (.06)	.78 (.17)	.78
Hearing	20	.46 (.06)	.89 (.09)	.88

*Note: "Transliteration" refers to English-based signing, yielding a two- (skill/preference) by-two (version received) design.

students enrolled in NTID might be less familiar with interpreting in STEM classrooms than are students who are registered in the other colleges of RIT, where interpreters work with deaf students and hearing faculty in the classroom. Repeating the analyses described earlier, separating cross-registered and NTID-registered students, however, reveals no reliable differences with regard to comprehension. The only demographic variable on which the two groups differed was the age at which they learned to sign, with the cross-registered students reporting having learned to sign later (mean = 5.4 years versus 2.9 years). Controlling for age of sign language acquisition, however, did not yield any reliable differences among the four groups of deaf students. Reanalyzing students' predictions concerning the number of questions they would get correct on the comprehension test, the cross-registered students (who rely on interpreters daily) showed a significant correlation with their actual test performance, ($r = .41$), whereas the NTID-registered students (who do not rely on interpreters for most classes) did not ($r = .29$).

General Discussion

The experiments described here address several important issues with regard to sign language interpreting. The original purpose of this

research concerned the effectiveness of interpreting versus transliteration for students' access to information in the classroom. In independent replications, the present experiments indicate that regardless of students' sign language skills and preferences, neither mode of interpreting is superior in terms of comprehension (see Mayer and Akamatsu [2000] for a similar finding with deaf adolescents). This conclusion can be made with confidence only for deaf individuals whose skills are within the range found among our samples of deaf university students, and a sample drawn from the community might yield different results. Nevertheless, the skills of the students we tested vary widely, from those who had been signing their whole lives, to others who learned to sign as adults.

Although the findings indicate a consistent lack of influence of students' reported sign language skills, they do not mean that the mode of interpreting cannot affect student performance. The comfort and ease of communication in a student's preferred mode might lead to increased motivation, participation, and learning in settings different from the technical and nontechnical lectures we used. It also may turn out that "live" interpreters lead to differences in comprehension not found with videotaped or remote interpreting, a possibility currently under investigation. For example, ASL may be more dependent on the use of three-dimensional space than English-based signing, thus biasing televised lectures against interpretation. Alternatively, ASL seems to involve more interaction between interpreters (monitoring) and deaf clients (feedback), again potentially creating a bias (cf. Fleischer 1975).

Such speculations aside, results from all three of the present experiments are consistent with those of Murphy and Fleischer (1977), who also found that interpretation and transliteration lead to comparable levels of performance in comprehending a nontechnical lecture. At the same time, such results raise the more basic question of how much of sign language interpreting deaf students really understand. Jacobs (1977) found that deaf students learn less in a mediated (interpreted) instructional setting than hearing peers do in the same classroom. Consistent with those findings and the findings of Livingston et al. (1994), comprehension performance following the nontechnical lecture in Experiments 1 and 2 was consistently around only 60 percent correct compared to hearing students' 87-percent-correct

performance. Deaf students in Experiment 3 did better on the comprehension test than they had in Experiments 1 and 2, but they still scored significantly below hearing peers (89 percent), even when prior content knowledge was controlled (see table 1).[4]

One possibility is that the deaf students perhaps understood more but that the use of a written test created a bias against them. However, Experiment 2 eliminated that alternative, as performance was essentially the same, regardless of whether comprehension was evaluated using a written test or a signed test and even when reading ability was statistically controlled.

Perhaps more telling is the fact that many deaf students apparently were unaware of the level of their comprehension. Many (if not most) students left the testing sessions of Experiments 1 and 2 remarking on how simple the questions were, with many suggesting that they got all of the answers correct. Relatively few deaf students had perfect comprehension scores, however, compared to the hearing students. In Experiment 3, in which a question on the test asked students to predict their performance, the hearing students' comprehension scores were significantly correlated with their expectations, whereas those of the deaf students as a whole were not (predictions of students who are more frequently involved with interpreters did, however, show a significant relation). Indeed, perhaps the fact that deaf students often have no way of knowing how much of an interpreted presentation they missed is the most disturbing aspect of these results (see Krinsky [1990] for similar findings with regard to vocabulary knowledge and Strassman [1997] with regard to reading). This issue is seldom raised publicly—although it is often discussed among interpreters—in part because it seems insensitive to suggest that deaf individuals may not be understanding high-quality interpreting. However, it is well recognized that there is considerable variability in exposure to and experience with sign language among deaf individuals; hence, differences in receptive sign skills should not be surprising (e.g., Napier 2002). Clearly, this issue needs careful exploration, especially when it concerns the education of deaf children and the frequently observed gaps in their academic and conceptual knowledge.

In summary, the present results clearly point to the need for more research into interpreting outcomes, including the effectiveness of

interpreting in various academic settings, and, more generally, into other aspects of interpreting and interpreter education. Without carefully controlled research concerning the student, interpreter, and contextual variables that likely influence comprehension and learning in interpreted settings, we cannot determine the best ways to train sign language interpreters. Of even greater concern is the possibility that mediated instruction via sign language interpreting, regardless of how accurate it is, may put deaf students at some risk for academic failure. At the university level, deaf students are often underprepared relative to hearing peers in terms of content and world knowledge (McEvoy, Marschark, and Nelson 1999; Stinson and Kluwin 2003). This situation requires awareness and appropriate modification of communication by both interpreters and instructors if those students are to have equal educational opportunities as well as equal access to information (Marschark, Lang, and Albertini 2002; Winston 1994).

Finally, it is noteworthy that all of the current and previous studies described here have involved interpreters who would minimally be described as well above average—but generally have been selected from the best available. What of the deaf student who is in a classroom with a less-experienced interpreter—one who has lesser skills—or those who are interpreting in science, mathematics, or some other field in which they lack content knowledge? Certainly, the observed academic difficulties of many deaf students do not lie solely, or even largely, in the quality of the sign language interpreting they receive. Nonetheless, the provision of appropriate interpreting in the classroom might ameliorate some of the academic challenges that deaf students face, and a number of relevant questions are still in need of answers.

Notes

1. Throughout this article, references to ASL and English are intended generically and refer to natural sign languages and the corresponding vernaculars of the hearing community. In keeping with existing terminology, the conversion of spoken English into ASL is referred to here as *interpreting*, and conversion of spoken language into various forms of English-based signing is referred to as *transliteration*.

2. The possible bias in the tenth (excluded) question from Experiment 1 was not changed, so as to allow comparisons between the two experiments.

3. It is noteworthy that without the necessary Bonferroni adjustment for multiple comparisons, there appeared to be significant, simple main effects. That result suggests that inconsistent findings in earlier studies involving multiple groups may have been due to inappropriate statistical methods.

4. The lack of a difference between hearing students' comprehension scores in Experiments 1 and 3 may have been due to ceiling effects, but this does not affect the interpretation of any of the results.

Preparation of this report was supported by grant #REC-0207394 from the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors wish to thank Michael Karchmer for his statistical advice and comments on a draft of this article. Correspondence may be addressed to Marc Marschark, National Sign Language Interpreting Project, National Technical Institute for the Deaf, 96 Lomb Memorial Drive, Rochester, NY 14623 (email: memrtl@rit.edu).

References

Baily, J., and S. Straub. 1992. Interpreting services. *American Speech-Language-Hearing Association* 34:51–53.

Caccamise, F., and R. Blasdell. 1977. Perception of Sentences under Oral-Manual Interpreted and Simultaneous Test Conditions. *American Annals of the Deaf* 122:414–21.

Fleischer, L. R. 1975. *Language Interpretation under Four Interpreting Conditions*. Ph.D. diss., Brigham Young University.

Harmer, L. M. 1999. Health Care Delivery and Deaf People: Practice, Problems, and Recommendations for Change. *Journal of Deaf Studies and Deaf Education* 4:73–110.

Harrington, F. 1999. Sign Language Interpreters and Access for Deaf Students to University Curricula: The Ideal and the Reality. In *The Critical Link 2: Interpreters in the Community,* ed. R. P. Roberts, S. E. Carr, D. Abraham, and A. Dufour. Amsterdam: John Benjamins.

Hatfield, N., F. Caccamise, and P. Siple. 1978. Deaf Students' Language Competency: A Bilingual Perspective. *American Annals of the Deaf* 123:847–51.

Jacobs, L. R. 1977. The Efficiency of Interpreting Input for Processing Lecture Information by Deaf College Students. *Journal of Rehabilitation of the Deaf* 11:10–14.

Johnson, R. E., S. Liddell, and C. Erting. 1989. *Unlocking the Curriculum: Principles for Achieving Access in Deaf Education*. Working paper 89-3, Gallaudet Research Institute, Gallaudet University, Washington, D.C.

Kluwin, T. N., and D. A. Stewart. 2001. Interpreting in Schools: A Look at Research. *Odyssey* (winter/spring):15–17. Washington D.C.: Gallaudet University.

Krinsky, S. G. 1990. The Feeling of Knowing in Deaf Adolescents. *American Annals of the Deaf* 135:389–95.

Lang, H. G. 2002. Higher Education for Deaf Students: Research Priorities in the New Millennium. *Journal of Deaf Studies and Deaf Education* 7:267–80.

Lang, H. G., B. G. McKee, and K. N. Conner. 1993. Characteristics of Effective Teachers: A Descriptive Study of Perceptions of Faculty and Deaf College Students. *American Annals of the Deaf* 138:252–59.

Leigh, G. R. 1995. *Teachers' Use of the Australasian Signed English System for Simultaneous Communication with Their Hearing-Impaired Students*. Ph.D. diss., Monash University, Melbourne, Australia.

Leigh, I. W., and R. Q. Pollard Jr. 2003. Mental Health and Deaf Adults. In *Oxford Handbook of Deaf Studies, Language, and Education,* ed. M. Marschark and P. E. Spencer, 203–16. New York: Oxford University Press.

Livingston, S., B. Singer, and T. Abramson. 1994. A Study to Determine the Effectiveness of Two Different Kinds of Interpreting. *Proceedings of the Tenth National Convention of the Conference of Interpreter Trainers—Mapping Our Course: A Collaborative Venture,* 175–97. CIT.

Marschark, M., H. Lang, and J. Albertini. 2002. *Educating Deaf Students: From Research to Practice*. New York: Oxford University Press.

Mayer, C., and C. T. Akamatsu. 1999. Bilingual-Bicultural Models of Literacy Education for Deaf Students: Considering the Claims. *Journal of Deaf Studies and Deaf Education* 4:1–8.

———. 2000. Deaf Children Creating Written Texts: Contributions of American Sign Language and Signed Forms of English. *American Annals of the Deaf* 145:394–410.

McEvoy, C., M. Marschark, and D. L. Nelson. 1999. Comparing the Mental Lexicons of Deaf and Hearing Individuals. *Journal of Educational Psychology* 91:1–9.

Monikowski, C., and E. A. Winston. 2003. Interpreters and Interpreter Education. In *Oxford Handbook of Deaf Studies, Language, and Education,* ed. M. Marschark and P. E. Spencer, 347–60. New York: Oxford University Press.

Murphy, H. J., and L. R. Fleischer. 1977. The Effects of Ameslan versus Siglish upon Test Scores. *Journal of Rehabilitation of the Deaf* 11:15–18.

Napier, J. 2002. University Interpreting: Linguistic Issues for Consideration. *Journal of Deaf Studies and Deaf Education* 7:281–301.

Napier, J., and R. Barker. 2004. University Interpreting: Expectations and Preferences of Deaf Students. *Journal of Deaf Studies and Deaf Education* 9:228–38.

Newell, W. 1978. A Study of the Ability of Day-Class Deaf Adolescents to Compare Factual Information Using Four Communication Modalities. *American Annals of the Deaf* 123:558–62.

Redden, M. R., C. A. Davis, and J.W. Brown. 1978. *Science for Handicapped Students in Higher Education: Barriers, Solutions, and Recommendations*. Washington, D.C.: American Association for the Advancement of Science.

Registry of Interpreters of the Deaf. 1995. *RID Membership Directory.* Silver Spring, Md.: RID.

Richardson, J. T. E., J. MacLeod-Gallinger, B. G. McKee, and G. L. Long. 1999. Approaches to Studying in Deaf and Hearing Students in Higher Education. *Journal of Deaf Studies and Deaf Education* 5:156–73.

Roy, C. B., ed. 2000. *Innovative Practices for Teaching Sign Language Interpreters*. Washington, D.C.: Gallaudet University Press.

Schein, J., D. Stewart, and B. Cartwright. 1998. *Sign Language Interpreting: Exploring Its Art and Science*. Boston: Allyn and Bacon.

Seal, B. C. 1998. *Best Practices in Educational Interpreting*. Boston: Allyn and Bacon.

Stinson, M. S., and T. Kluwin. 2003. Educational Consequences of Alternative School Placements. In *Oxford Handbook of Deaf Studies, Language, and Education,* ed. M. Marschark and P. E. Spencer, 52–64. New York: Oxford University Press.

Stewart, D. A., and T. N. Kluwin. 1996. The Gap between Guidelines, Practice, and Knowledge in Interpreting Services for Deaf Students. *Journal of Deaf Studies and Deaf Education* 1:29–39.

Strassman, B. 1997. Metacognition and Reading in Children Who Are Deaf: A Review of the Research. *Journal of Deaf Studies and Deaf Education* 2:140–49.

Traxler, C. B. 2000. Measuring Up to Performance Standards in Reading and Mathematics: Achievement of Selected Deaf and Hard-of-Hearing Students in the National Norming of the Ninth Edition Stanford Achievement Test. *Journal of Deaf Studies and Deaf Education* 5:337–48.

Winston, E. A. 1994. An Interpreted Education: Inclusion or Exclusion? In *Implications and Complications for Deaf Students of the Full Inclusion Movement,* ed. R. C. Johnson and O. P. Cohen, ed., 55–62. Gallaudet Research Institute Occasional Paper 94-2. Washington, D.C.: Gallaudet University.

Appendix
Communication Questionnaire, Experiment 3

1. How much of the interpreter's fingerspelling did you understand?

All of it Some of it None of it

2. How much of the information in the lecture was new to you?

   All of it Some of it None of it

3. Please circle all that are true:
   a. I understood the interpreter very well (signs and fingerspelling).
   b. I understood the signs but not the fingerspelling.
   c. I did not understand the interpreter.
   d. The interpreter signed clearly, but I did not understand the lecture.

4. The interpreter signed:

   Too fast Too slow At a comfortable pace to follow

5. I would understand this material *best* from (choose ONE):
   a. An interpreter
   b. A teacher signing
   c. Reading about it
   d. Captions
   e. C-print

6. Overall, how would you rate this interpreter?

   Excellent Okay Poor

7. How do you prefer to communicate most of the time? (please circle only one):
   a. Sign alone
   b. Speech alone
   c. Sign and speech together (simultaneous communication)
   d. Other _____ (please specify)

8. Please rate your skills in understanding simultaneous communication (speech and sign together). I understand (circle one):

   Everything 5 4 3 2 1 Nothing

9. a. Please rate your skill in *producing* ASL (circle one):

Excellent 5 4 3 2 1 No skill

9. b. Please rate your skill in *understanding* ASL (circle one):

Excellent 5 4 3 2 1 No skill

9. c. Please rate your skill in *producing* signed English:

Excellent 5 4 3 2 1 No skill

9. d. Please rate your skill in *understanding* signed English (no voice):

Excellent 5 4 3 2 1 No skill

10. Age you began to learn sign language:
    a. Since birth
    b. _____ years old
    c. Do not know sign language

11. Do you use a hearing aid? Yes ____ No ____

12. Do you use a cochlear implant? Yes ____ No ____

13. Do you use another kind of Assistive Listening Device?
    Yes ____ No ____

14. Do you use a spoken language other than English with your family?
    Yes ____ No ____ If yes, please specify _____

Please *circle one number* to answer each of the following questions:

15. *Overall,* I prefer to use

Sign language 5 4 3 2 1 Spoken language

16. *Overall,* I prefer to use

ASL 5 4 3 2 1 English-based signing