Enhancing Deaf Students’ Learning from Sign Language and Text: Metacognition, Modality, and the Effectiveness of Content Scaffolding

Georgianna Borgna1, Carol Convertino1, Marc Marschark*1,2, Carolyn Morrison1, Kathleen Rizzolo1

1National Technical Institute for the Deaf—Rochester Institute of Technology
2University of Edinburgh

Received June 3, 2010; revisions received July 28, 2010; accepted July 29, 2010

Four experiments, each building on the results of the previous ones, explored the effects of several manipulations on learning and the accuracy of metacognitive judgments among deaf and hard-of-hearing (DHH) students. Experiment 1 examined learning and metacognitive accuracy from classroom lectures with or without prior scaffolding in the form of a description of main points and concepts. Experiment 2 compared the benefits of scaffolding when material was read versus when it was presented as a lecture signed for DHH students and spoken for hearing students. Experiment 3 compared scaffolding provided in the form of main points versus vocabulary, and Experiment 4 examined effects of material familiarity and a delay between study and test. Results indicated that although all students had a tendency to overestimate their performance, hearing students learned more and were more accurate in their metacognitive judgments than DHH students. Content familiarity improved the accuracy of metacognitive judgments by both DHH and hearing students, but the delay manipulation was effective only for hearing students. Consistent with other recent findings, DHH students learned as much from reading as they did from signed instruction. Differences between DHH and hearing students may indicate the need for explicit instruction for DHH students in academically relevant skills acquired incidentally by hearing students.

This article explores metacognition and learning among deaf and hard-of-hearing (DHH) college students. Most research addressing relations of metacognition and learning in this population has examined reading (e.g., Krinsky, 1990; Strassman, 1997; Walker, Munro, & Rickards, 1998), an area of particular interest and concern among researchers and educators of DHH students of all ages. Although some DHH students read as well as hearing age-mates, most lag behind in reading subdomains of vocabulary, syntax, inferencing, and comprehension more generally (Brown & Brewer, 1996; Traxler, 2000; Trezek, Wang, & Paul, in press). Marschark et al. (2009) argued that the available evidence indicates that the challenges observed in DHH students’ reading are not really about reading per se. For that and other reasons to be made clear below, this investigation examined learning both from text and from classroom lectures.

Although fully supporting the goals of equality and the right to equal opportunities, several investigators over the past decade have demonstrated that by virtue of differing backgrounds, language experiences, and cognitive/brain processes, DHH individuals are not just hearing individuals who cannot hear (see chapters in Marschark & Hauser, 2008, for reviews). Some of the cognitive differences observed between DHH and hearing individuals may be advantageous (e.g., enhanced sensitivity to peripheral visual signals and face discrimination abilities). Others, however, are likely to impede comprehension and learning in educational settings. DHH students, for example, generally have been found less likely than hearing

*Correspondence should be sent to Marc Marschark, National Technical Institute for the Deaf, 52 Lomb Memorial Drive, Rochester, NY 14623 (e-mail: marc.marschark@rit.edu).
peers to utilize prior knowledge in acquiring new information (see Hauser, Lukomski, & Hillman, 2008; Marschark, Convertino, & LaRock, 2006). They also have been shown to be less accurate than hearing peers in knowing what they know. Krinsky (1990), for example, examined DHH high school students’ awareness of their own vocabulary knowledge. Students provided judgments of their feeling of knowing for words they had missed on a vocabulary test. The accuracy of their predictions was determined by having them match the words to pictures. Relative to hearing students of the same reading ages, DHH students were unable to accurately predict their performance (i.e., word knowledge). Those reading results parallel findings of Marschark, Sapere, Convertino, Seewagen, and Maltzen (2004) and Marschark, Sapere, Convertino, and Seewagen (2005) involving DHH students’ sign language comprehension. DHH and hearing college students in their experiments were asked to predict their performance on content-specific tests following signed classroom lectures. DHH students consistently overestimated their learning relative to hearing classmates, and, as in the Krinsky study, predicted and obtained scores were not significantly related.

It is unclear to what extent the lesser metacognitive abilities exhibited by DHH students in reading and problem solving are a function of a lack of academic skills training versus the acknowledged variability or impoverishment in early language, cognitive, and educational environments. The importance of the mutual influence among these domains, however, should not be underestimated. On the basis of an extensive review of the empirical literature, Spencer and Marschark (2010) noted that cognitive growth will be affected if a DHH (or any other) student “lacks sufficient language sophistication to allow ‘thinking about’ learning (i.e., metacognition), organizing and coding of information to support memory, inferencing, and the drawing of logical conclusions based on understanding nuance” (p. 50). Accordingly, research involving (hearing) students who are higher or lower achievers, good or poor readers, and better or worse problem solvers consistently has indicated that better students have better metacognitive skills and vice versa. In particular, despite the usual focus on reading, metacognitive skills are an essential part of all-language comprehension (and learning through language) regardless of modality.

Sinkavich (1995) examined the relation between hearing students’ judgments of their performance and actual test performance. Defining good and poor students as those scoring in the top and bottom thirds on examinations, he found that the good students were significantly more accurate in predicting their performance than the poor students. “This result indicates that good students are better able to self-regulate their own learning” (p. 83). Maki, Shields, Wheeler, and Zacchilli (2005) also examined metacognition in hearing college students’ learning from text. They found that students with lower verbal abilities tended to overestimate their subsequent test performance, whereas those with higher verbal abilities tended to underestimate their performance. This pattern of results has been referred to as the unskilled and unaware effect (Kruger & Dunning, 1999) because less content knowledge or poor language comprehension skills may leave an individual with the “double burden” of poor comprehension and less awareness of it. Unfortunately, many DHH students appear to carry that double burden.

Gibbs (1989) observed that good readers, both DHH and hearing, utilize metacognitive strategies to monitor and maintain comprehension (e.g., “look-back” when uncertain about whether a passage was correctly understood). DHH students, however, tend to use fewer such strategies, less often, and generally are less accurate in their metacognitive judgments and self-monitoring than hearing peers. Strassman (1997) and Marschark, Lang, and Albertini (2002, chapter 8) argued that the situation is caused in part by the way that DHH students are taught and, in particular, the use of simplified reading materials. Emphasis on reading subskills, memorizing vocabulary words, and answering teacher questions takes away from the reading of authentic texts for meaning and may lead students to adopt relatively superficial comprehension criteria while failing to acquire the metacognitive strategies necessary for fluent reading.

Similar arguments have been made with regard to DHH students’ performance in mathematics and science. Pagliaro and Ansell (2002) concluded that one
reason why DHH students perform so poorly on mathematics story problems is that opportunities to engage in such problem solving rarely occur in DHH classes. Less than 20% of the 36 first- through third-grade teachers they surveyed presented story problems daily. Teachers reportedly believed that such problems, whether presented in sign language, spoken language, both, or as text, are too difficult for DHH children until basic math skills and reading skills are achieved. Pagliaro and Ansell argued that such problems should be used from the earliest grades to engage DHH children in mathematical thinking, problem-solving processes, and related metacognition (p. 116).

Yore (2000) recommended embedding structured writing activities within the teaching of science in order to guide students’ thinking and encourage active evaluation of their own knowledge. He suggested that following a concrete, hands-on activity, writing about it would support the integration of ideas and help students address relationships they have discovered. Yore emphasized the importance of cognitive and metacognitive skills for science learning, arguing that effective reading and writing in science require “conceptual background; knowledge about science text and science reading; declarations, procedures, and conditions of reading strategies; and executive control to set purpose, monitor progress, and adjust actions” (p. 110). Underpinning all this is the student’s awareness of what they already know and strategies for applying prior knowledge in new contexts, that is, metacognition.

Aims of this Article

This investigation was prompted by a convergence of evidence indicating that DHH students tend not to apply knowledge they have or are given in contexts where it would be helpful and frequently do not accurately evaluate their ongoing comprehension and learning. This is not to place the entire responsibility for that situation on the student. A variety of studies has found DHH students to be relatively passive readers and problem solvers, rather than actively engaging in metacognitively guided strategies, unless directed by the teacher; and teachers often foster such instrumental dependence in their DHH students (Marschark et al., 2002, p. 198). Such results are obtained with both verbal (Schirmer, 2003; Schirmer, Bailey, & Lockman, 2004) and nonverbal (Marschark & Everhart, 1999; Mousley & Kelly, 1998; Senior, 2004) materials, indicating that the issue goes beyond reading.

On the anecdotal side, we see DHH college students surprised at receiving poor final grades in courses despite having failed multiple exams (and/or not attended class) or turning in revised homework assignments with errors identical to those pointed out in the originals. In previous studies, we frequently have seen DHH students respond incorrectly on content-specific pretests and then, after seeing a lecture on the topic, repeat the same errors on the identical questions. These examples are not necessarily specific to DHH students, but they do seem to occur more frequently than among hearing students. In any case, such anecdotes emphasize the joint functioning of learning and active self-monitoring during related activities, a point of intersection that clearly is in need of improvement in the education of DHH students.

Following on the findings of Marschark, Sapere, et al. (2004, 2005) that DHH students were less accurate than hearing students in predicting test performance following classroom lectures, this study examined both learning via text and learning via sign language. The specific aim was to improve knowledge acquisition by providing support for DHH students’ ongoing comprehension and promoting greater metacognitive awareness. Thiede and Anderson (2003) and Rawson and Kintsch (2002) demonstrated that greater comprehension leads to more accurate metacognitive judgments. Those findings are relevant here at two different levels: (a) DHH students generally were expected to be less accurate in their metacognitive judgments than hearing students regardless of the modality of presentation and (b) increasing comprehension through the provision of scaffolding was expected to improve metacognitive accuracy and, as a result of enhanced understanding and self-monitoring, to improve learning. The latter effects were predicted to be larger for DHH than hearing students, assuming that the latter group is more likely to deploy metacognitive strategies spontaneously and thus is less in need of supporting materials.
The provision of scaffolding materials to DHH students is a common practice, but one that appears to be based more on intuition than empirical evidence. Stewart and Kluwin (2001) suggested that “[w]hen a teacher engages in scaffolding, she or he is providing deaf students with whatever assistance they need to do the task at hand” (p. 85). Schirmer (2000) also claimed scaffolding to be a particular benefit to DHH students suggesting that “As the teacher questions, expands, and re-states in response to the child’s written comments the teacher is in essence providing a scaffold of language learning for the deaf child” (p. 177). Yet, the citations those authors provided were all drawn from literature involving hearing students. Still to be determined is what makes for effective scaffolding for DHH students, who vary widely in their language, cognitive, and academic skills.

Experiment 1

This experiment was designed to provide empirical evidence concerning the use of scaffolding in classes involving DHH students. DHH and hearing students received scaffolding in the form of summary statements that included the key themes and concepts of two classroom lectures. On the basis of previous studies, it was expected that hearing students would surpass DHH students in learning, overall, and in the accuracy of their metacognitive judgments (Marschark, Sapere, et al., 2004, 2005). Nevertheless, it was hoped that the provision of explicit scaffolding—and explanation of its potential benefit (Fox, 1994)—would lead to significant improvements in DHH students’ learning and metacognitive accuracy. The benefits of scaffolding were expected to be larger for DHH than hearing students in part because there would be greater room for improvement and also because hearing students would be more likely to engage in integrative comprehension strategies spontaneously (Marschark, Convertino, et al., 2006).

Method

Participants. Twenty DHH students and 20 hearing students participated. All were enrolled at Rochester Institute of Technology (RIT), recruited via campus advertising, and paid for their participation. RIT includes the National Technical Institute for the Deaf (NTID), but participants were recruited from across all RIT colleges. Because DHH students’ hearing thresholds generally do not predict learning or achievement, that information was not solicited (see Note 1). Incoming DHH students at RIT, however, are expected to have unaided hearing thresholds of 70 dB or higher in the better ear. In fact, the average of about 94 dB has not changed over the past 20 years, ranging from 92.6 to 96.5 dB since 1998 (NTID, 2009).

It may be noteworthy that RIT records indicate the vast majority of DHH students, both those in NTID and those in the other colleges of RIT, to have attended mainstream schools for all or most their prior academic careers. In their large-scale study, however, Convertino, Marschark, Sapere, Sarchet, and Zupan (2009) found that enrollment in NTID programs (all DHH) versus RIT programs (approximately 5% DHH) did not significantly predict either academic achievement (college entrance test scores) or classroom learning when other factors were controlled.

Materials and procedure. Two hearing members of the RIT faculty were videotaped giving introductory-level lectures, approximately 15 min in length; both were blind to the purpose of the study. One lecture was on soil mechanics and the other was on visual perception. Digital recordings of the lectures were made in a small studio with lighting designed to avoid shadow. Interpreted versions of the lectures were provided by an RIT interpreter with more than 20 years of experience, recommended by his departmental manager for his outstanding skill in RIT classrooms. Their interpretations were rendered in the most common mode used in RIT classrooms: signing primarily with English word order but with characteristics of American sign language (ASL; see Marschark, Sapere, et al., 2004, for an empirical comparison to ASL). As would be the case in a typical classroom, the interpreters did not see the materials in advance; they were videotaped in the same manner as the instructors.

Naturalistic tests of lecture content were developed by the investigators in collaboration with the instructors. For each lecture, there were 12 multiple-choice
questions with four plausible answer options, only one of which was correct. The investigators also created summary statements for each lecture (i.e., scaffolds) that, taken together, outlined what the lecture was about and highlighted important concepts without providing answers to any of the multiple-choice questions. For the soil mechanics lecture, the scaffolding statement was “In this lecture, you will learn about the study of soil, what soil is made of, and what the study of soil is called. The meaning of compressibility and grain size distribution will be given, and the proper terminology for the spaces between the soil particles where liquid and gases are found.” For the visual perception lecture, the scaffolding statement was “The professor will explain the function of each part of the eye enabling us to adjust to light, dark, and form images. You will hear terms like, fovea and resolution, and will hear an explanation of how the eye focuses. You also will learn the name of the parts of the eye that detect black, white, and color.”

Testing was conducted in small groups, mixed with regard to hearing status as in RIT classes. In each group, students saw both lectures; lecture order was balanced across groups. The lectures and their interpretations were presented side-by-side via life-sized video projection in a mock classroom (see Marschark, Pelz, et al., 2005, for evidence indicating no difference in learning between video projection and live presentation with these materials). Testing was conducted by a senior sign language interpreter who used both spoken language and sign language because of the mixed groups.

The first lecture was presented without scaffolding. Immediately afterward, students were asked to write down what they thought were the five most important points of the lecture. This offered students the opportunity to integrate and consolidate the information provided (Dowaliby & Lang, 1999; Thiede, Anderson, & Therriault, 2003; Wittrock & Alessandrini, 1990) and gave the investigators an alternative means of evaluating student learning. This generation task was followed by the multiple-choice test, after which students were asked to indicate for each question whether they thought they had gotten it right or wrong (Maki & Serra, 1992). Prior to the second lecture, the nature of the scaffolding was explained, and students were told that it would help them to understand the lecture. They then were given the scaffolding in both spoken and sign language. The procedure then followed the same sequence as the first lecture.

Results and Discussion

Unless otherwise noted, all effects reported here and in subsequent experiments were significant at the .05 level, and only those effects are reported.

Multiple-choice tests. Because of the incomplete blocks design (i.e., students saw both lectures but under different conditions), multiple-choice test scores were analyzed separately for each lecture using a 2 (hearing status: DHH vs. hearing) x 2 (scaffolding or no scaffolding) analysis of variance. Significant main effects of hearing status were obtained for both the soil mechanics lecture, F(1, 36) = 32.47, MSE = 218.03, and the visual perception lecture, F(1, 36) = 45.80, MSE = 203.31 (see Table 1). Although there was not a significant main effect of scaffolding for either lecture, scaffolding led to slightly better performance for both groups with both lectures. A priori t tests indicated that the scaffolding of lectures was a significant benefit for DHH students, t(19) = 2.32, but not hearing students, t(19) = 1.31. DHH students gained 20% from prior scaffolding compared to only 6% for the hearing students. This difference may partially reflect a ceiling effect for the hearing students, although their average scores of 81% without scaffolding and 86% with scaffolding suggest that is unlikely. Alternatively, it may be that providing the scaffolding was superfluous for the hearing students, redundant with whatever learning strategies they already were applying. Even with scaffolding, however, DHH students’ mean test score (60%) did not reach that of hearing students’ without scaffolding (81%). DHH students’ scores averaged approximately 66% of those obtained by hearing students, consistent with previous studies (e.g., Marschark, Sapere, et al., 2004), but the trend in the data suggests that some form of scaffolding might be particularly helpful for DHH students (see below).

Main ideas. For the purposes of scoring the main ideas generated by the students, four investigators
watched the lectures with the goal of identifying the five most important points in each. The overlap in their lists consisted of five points, which subsequently were used in scoring students’ five generated ideas. The latter were scored using a 4-point rubric: correct and complete (3), mostly correct and mostly complete (2), close but no cigar (1), and no credit (0), yielding a maximum possible score of 15 for each lecture. Students’ total main idea scores were analyzed as above. Consistent with their performance on the multiple-choice test, hearing students were better at reproducing the main ideas of the lectures than their DHH peers for both the soil mechanics lecture, $F(1, 36) = 21.64, MSE = 8.74$, with no other significant effects (see Table 1).

Metacognitive judgments. Previous studies have indicated that different individuals overestimate and underestimate their performance in different areas, at least partly as a function of their knowledge in that area (less and more, respectively). Metacognitive judgment scores for participants’ underestimations and overestimations were calculated by subtracting the accuracy of the prediction from the accuracy of the response ($correct = 1$, $incorrect = 0$, in both cases) for each question. The summed frequencies of underestimations (predicting that responses were incorrect when they were correct) and overestimations (predicting that responses were correct when they were not) were analyzed using a 2 (DHH vs. hearing) × 2 (scaffold vs. no scaffold) × 2 (prediction accuracy: underestimations vs. overestimations) design in which the last factor was within subjects. A separate analysis of variance was conducted for each lecture due to the incomplete blocks design, but both analyses yielded essentially the same results. For both the soil mechanics and the visual perception lectures, respectively, significant main effects of hearing status, $F(1, 36) = 19.66$, $MSE = 1.65/F(1, 36) = 34.35$, $MSE = 1.35$, and prediction accuracy, $F(1, 36) = 7.92$, $MSE = 2.16/F(1, 36) = 75.85$, $MSE = 1.62$, were obtained, and there was an interaction of hearing status and prediction accuracy, $F(1, 36) = 12.01/F(1, 36) = 23.11$. The main effects reflect the findings that the hearing students were more accurate in their metacognitive judgments than the DHH students and that both DHH and hearing participants generally were more likely to overestimate rather than underestimate their accuracy. The interaction resulted from the fact that DHH students were much more likely to overestimate rather than underestimate their accuracy relative to their hearing peers (see Table 1).

To summarize, performance measured both by postlecture multiple-choice tests and generation of the main ideas of the lectures indicated that hearing students learned more than did DHH students. The present results thus replicate earlier findings indicating both that DHH students in mainstream college classrooms learn less, on average, than their hearing peers.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Means and standard errors for test scores (%), generated main ideas (score of 15 possible), and metacognitive judgments (frequencies): Experiment 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deaf Mean</td>
</tr>
<tr>
<td>Multiple-choice test</td>
<td></td>
</tr>
<tr>
<td>Soil mechanics</td>
<td></td>
</tr>
<tr>
<td>Without scaffold</td>
<td>0.56</td>
</tr>
<tr>
<td>With scaffold</td>
<td>0.63</td>
</tr>
<tr>
<td>Visual perception</td>
<td></td>
</tr>
<tr>
<td>Without scaffold</td>
<td>0.45</td>
</tr>
<tr>
<td>With scaffold</td>
<td>0.58</td>
</tr>
<tr>
<td>Main ideas</td>
<td></td>
</tr>
<tr>
<td>Soil mechanics</td>
<td></td>
</tr>
<tr>
<td>Without scaffold</td>
<td>4.00</td>
</tr>
<tr>
<td>With scaffold</td>
<td>4.18</td>
</tr>
<tr>
<td>Visual perception</td>
<td></td>
</tr>
<tr>
<td>Without scaffold</td>
<td>1.54</td>
</tr>
<tr>
<td>With scaffold</td>
<td>3.33</td>
</tr>
<tr>
<td>Metacognitive judgments</td>
<td></td>
</tr>
<tr>
<td>Soil mechanics</td>
<td></td>
</tr>
<tr>
<td>Without scaffold</td>
<td></td>
</tr>
<tr>
<td>Underestimation</td>
<td>1.80</td>
</tr>
<tr>
<td>Overestimation</td>
<td>2.90</td>
</tr>
<tr>
<td>With scaffold</td>
<td></td>
</tr>
<tr>
<td>Underestimation</td>
<td>0.90</td>
</tr>
<tr>
<td>Overestimation</td>
<td>3.20</td>
</tr>
<tr>
<td>Visual perception</td>
<td></td>
</tr>
<tr>
<td>Without scaffold</td>
<td></td>
</tr>
<tr>
<td>Underestimation</td>
<td>0.90</td>
</tr>
<tr>
<td>Overestimation</td>
<td>4.30</td>
</tr>
<tr>
<td>With scaffold</td>
<td></td>
</tr>
<tr>
<td>Underestimation</td>
<td>0.90</td>
</tr>
<tr>
<td>Overestimation</td>
<td>4.60</td>
</tr>
</tbody>
</table>
peers and think that they are learning more than they really are (Marschark, Sapere, et al., 2004, 2005). To the extent that the locus of these findings lies either in DHH students’ coming into the classroom with less content knowledge than their hearing peers \(^3\) (see Marschark, Pelz, et al., 2005) or lesser language comprehension skills (Marschark et al., 2009), these results could reflect an unskilled and unaware effect. That is, students may not recognize when they are not fully comprehending the content (e.g., they understand the individual signs/words but not the relations among them) and therefore do not engage in the cognitive activity necessary to assimilate it fully. Thiede and Anderson (2003) similarly argued that the less well material is understood, the less able an individual is to know how well they have understood it (Kruger & Dunning, 1999). Results from the soil mechanics lecture were consistent with that prediction (scaffolding led to more accurate metacognitive judgments), but those from the visual perception lecture were inconsistent. At this point, it is worth noting that prior to analyzing any of the data, the investigators agreed that the visual perception was the more difficult of the two lectures. Examination of Table 1 reveals that in addition to yielding the expected pattern of metacognitive judgments, the soil mechanics lecture yielded slightly better scores on all dimensions examined. This issue will be reconsidered later.

The present results suggest that providing DHH students with prior information concerning the content and structure of a lecture can enhance their learning to some (limited) extent. Still to be determined is whether more effective scaffolding can be developed, perhaps providing a structure more congruent with the knowledge and cognitive processes that DHH students’ bring to the classroom. That is, although the scaffold provided in this experiment increased DHH students’ test scores, it may be that the summary points provided by the investigators did not coincide with DHH students’ knowledge or its organization and thus were less than optimally effective. Previous studies have demonstrated that students’ background knowledge plays an essential role in ongoing learning (Rawson & Kintsch, 2002) but also that there are significant quantitative and qualitative differences between DHH and hearing students’ concept knowledge (Marschark, Convertino, McEvoy, & Masteller, 2004; McEvoy, Marschark, & Nelson, 1999). With regard to generation of main ideas, a related possibility is that what DHH students saw as important in the lecture was different than that of both the hearing students and the (hearing) instructors and investigators.

A third factor potentially affecting these results relates to the cognitive processes involved in utilizing a scaffold. Thiede et al. (2003) demonstrated that the generation of content-descriptive keywords following reading of a passage significantly enhanced accuracy of related metacognitive judgments, and Thiede and Anderson (2003) found similar effects following summarization of passage. It may be that self-generation of organizing information is an essential component of scaffolding effects on judgment accuracy. Finally, although providing content outlines for students is a common teaching strategy, it may be that the college students, and DHH college students in particular, do not immediately recognize the organizational potential of such materials. Experiments 2 and 3 considered the viability of these alternatives.

**Experiment 2**

Experiment 1 explored learning through the presentation of classroom lectures followed by content-specific tests. Understanding what it is that DHH students take away from reading is also relevant to increasing how much they take away from the classroom in part because prior content and world knowledge derived from reading help them to comprehend and acquire new information (Rawson & Kintsch, 2002). The reverse is also true, as what students learn in class provides the framework for understanding what they read. DHH students’ reading comprehension has been studied far more than their classroom learning, however, and there appears an implicit assumption that their learning from text and learning from classroom instruction are fundamentally different activities.

Contrary to naive expectations, recent studies have failed to find quantitative differences in how much DHH college students learn from signed lectures and real-time text in the classroom (Marschark, Leigh, et al., 2006; Stinson, Elliot, Kelly, & Liu, 2009),
although Stinson et al. (2009) obtained contradictory evidence with regard to younger students. Similarly, Marschark et al. (2009) found that DHH students understood as much or more information obtained from reading a passage as they did from seeing it signed, a comparison apparently not made previously. In all those situations, DHH students have been found to learn 65%–80% of what their hearing peers do, regardless of the modality of presentation. Whether this “glass ceiling effect” is related to less than optimal language comprehension or metacomprehension strategies (Marschark & Wauters, 2008), cognitive skills that are not fine-tuned during the school years (Hauser et al., 2008), or some other factor remains to be determined.

Experiment 2 extended Experiment 1 and explored these issues by examining DHH and hearing students’ learning from both a classroom lecture and text. In addition to receiving a postlecture multiple-choice test, students were again asked to generate the important points of the material. The “correctness” of main ideas generated was considered both in terms of what content area instructors identified as the important points in the material and what other students, both DHH and hearing, indicated was most important. Scoring the generated main ideas against “objective” standards of instructors as well as the perceptions of their peers allowed us to examine qualitative as well is quantitative differences in what DHH and hearing students learn from text and from signed classroom presentations.

Method

Participants. Twenty DHH and 20 hearing students were recruited from the same student population in Experiment 1 and were paid for their participation.

Materials. Two passages were drawn from college-level biology textbooks, one on Gregor Mendel’s work in genetics (770 words) and one on circadian rhythms (859 words). Spoken versions of the passages were provided by an RIT science instructor. Although being videotaped, he read each via a video projection system, Teleprompter fashion, and spoke them as naturally as possible as though in the classroom. The signed versions were provided by a (hearing) NTID faculty member, who is a native user of ASL and a certified interpreter. After familiarizing himself with the two passages, he interpreted the material from the spoken lecture.

Multiple-choice tests were constructed as in the previous experiment. The “objective” main ideas were obtained by having three biology instructors generate up to the four most important points in each passage. Their overlap resulted in five main ideas for each passage to be used in scoring. Student-oriented main ideas were obtained by having 12 DHH and 10 hearing students participating in a preliminary study read the passages and highlight up to the four most important points in each passage.

Not surprisingly, student-identified main ideas were less consistent than instructor-identified main ideas. A criterion of 50% agreement on main points (separate for DHH and hearing students) resulted in three consensus ideas for both DHH and hearing students for the circadian rhythms passage as well as hearing students on the Mendel passage; there were also three consensus ideas for DHH students on the Mendel passage using a 40% criterion. DHH and hearing students identified only one main idea in common among the top three in the circadian rhythms passage; they agreed on two of three in the Mendel passage. For the purposes of student–student comparisons, main ideas generated in this experiment therefore were compared to the three most frequently identified in the preliminary study. For both DHH and hearing student with both passages, all three of those main points fell among the five identified by the instructors.

Procedure. Participants were tested in small groups, separate for DHH and hearing students, by two senior sign language interpreters who used spoken language and/or sign language as appropriate. Students first read one passage or they saw it signed (DHH students) or spoken (hearing students) and then wrote down what they considered the four most important ideas. Students then completed the appropriate multiple-choice test. The procedure was repeated for the second passage, which was presented in the mode opposite the first passage. Passage order and presentation mode were balanced over testing sessions.
Results and Discussion

Because of the incomplete blocks design, data for the two passages were analyzed separately using separate 2 (hearing status) × 2 (presentation mode) analyses of variance in which both factors were between subjects.

**Multiple-choice tests.** Analysis of test scores associated with the Mendel passage yielded both a significant effect of hearing status, $F(1, 36) = 15.98, MSE = 0.018$, as hearing students outscored DHH students and an effect of mode of presentation, $F(1, 36) = 5.90, MSE = 0.018$, as both DHH and hearing students scored higher when the passage was read than when it was signed or spoken (see Table 2). A similar analysis using scores from the circadian rhythms test yielded only a main effect of hearing status, $F(1, 36) = 22.10, MSE = 0.017$, again in favor of hearing students. As shown in Table 2, both groups did equally well on the circadian rhythms passage regardless of whether it was read or signed/spoken. DHH students performed considerably better in this experiment than an Experiment 1, although their scores still averaged only about 78% of those obtained by hearing students comparable to that obtained in earlier studies (e.g., Marschark, Sapere, et al., 2005).

**Main ideas.** Main ideas generated by the participants first were scored using the instructor-designated criteria and the same procedure as in Experiment 1. Although differences in scores followed the same pattern as in Experiment 1, no significant effects were observed in an analysis of variance using the same design as the multiple-choice test. Scores were quite low, however, as the mean scores for DHH students and those for hearing students were only 2.65 and 3.15, respectively, for the Mendel passage and 2.30 and 3.25 for the circadian rhythms passage out of a possible 15 points for each (see Table 2). The lack of significant differences in the above analysis thus may have been due to a floor effect rather than any particular benefit of either generating main ideas or the mode of presentation.

Main ideas scored against student criteria were analyzed separately for each passage using 2 (hearing status) × 2 (presentation mode) × 2 (scoring criteria: DHH or hearing student-identified main points) analyses of variance in which the last factor was within subjects. The Mendel passage yielded a main effect of hearing status, $F(1, 36) = 8.36, MSE = 0.43$, as hearing students outscored DHH students, and a presentation mode by scoring criteria interaction, $F(1, 36) = 5.24, MSE = 0.019$, as scores were higher when a passage was signed/spoken using the DHH students’ scoring criteria but higher when the passage was read using the hearing students’ scoring criteria. Scoring the circadian rhythms passage against student criteria yielded only a main effect of presentation mode, $F(1, 36) = 6.37, MSE = 0.50$, as higher scores were obtained when the passage was signed/spoken. Scoring according to student criteria resulted in

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Means and standard errors for test scores (%), generated main ideas (score of 15 possible), and metacognitive judgments (frequencies): Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deaf</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Multiple-choice test Mendel</td>
<td></td>
</tr>
<tr>
<td>Signed/spoken</td>
<td>0.65</td>
</tr>
<tr>
<td>Read</td>
<td>0.80</td>
</tr>
<tr>
<td>Circadian rhythms</td>
<td></td>
</tr>
<tr>
<td>Signed/spoken</td>
<td>0.62</td>
</tr>
<tr>
<td>Read</td>
<td>0.60</td>
</tr>
<tr>
<td>Main ideas (instructor generated) Mendel</td>
<td></td>
</tr>
<tr>
<td>Signed/spoken</td>
<td>2.10</td>
</tr>
<tr>
<td>Read</td>
<td>3.20</td>
</tr>
<tr>
<td>Circadian rhythms</td>
<td></td>
</tr>
<tr>
<td>Signed/spoken</td>
<td>2.60</td>
</tr>
<tr>
<td>Read</td>
<td>2.00</td>
</tr>
<tr>
<td>Metacognitive judgments Mendel</td>
<td></td>
</tr>
<tr>
<td>Signed/spoken</td>
<td>Underestimation</td>
</tr>
<tr>
<td>Overestimation</td>
<td>2.10</td>
</tr>
<tr>
<td>Read</td>
<td>Underestimation</td>
</tr>
<tr>
<td>Overestimation</td>
<td>3.90</td>
</tr>
<tr>
<td>Circadian rhythms</td>
<td>Signed/spoken</td>
</tr>
<tr>
<td>Overestimation</td>
<td>3.40</td>
</tr>
<tr>
<td>Read</td>
<td>Underestimation</td>
</tr>
<tr>
<td>Overestimation</td>
<td>3.30</td>
</tr>
</tbody>
</table>
extremely low scores for both passages (less than 1.0 of a possible 9.0 in all cells) reflecting the greater idiosyncrasy in DHH and hearing students’ selection of main ideas—essentially an unskilled and unaware effect relative to instructor criteria. This finding raises the issue of the value of “peer tutoring” in the classroom.

A number of educators have suggested that one benefit to congregated settings for DHH students is the opportunity for students to teach each other and clarify when instruction is not fully/correctly understood (e.g., Herring-Harrison, Gardner, & Lovelace, 2007; see Spencer & Marschark, 2010, for a review). Fox (1994), for example, utilized a group collaboration program supported by instructor-provided discussion questions in seeking to foster greater understanding of classroom readings among DHH college students. He found that “[s]omewhat more participation in discussion did result, but not greater understanding of the texts. Students still seemed just as confused in their questioning and did just as poorly on the tests” (p. 508). The findings described earlier indicating that DHH students understand less than they think they do in the classroom suggest that such activities need to be carefully monitored and empirically validated. In any case, scoring in this experiment based on what peers thought was important in the materials clearly underrepresented what DHH and hearing students learned, and the method was not used in subsequent experiments.

It is tempting to conclude that the active generation of main ideas in this experiment is responsible for the lack of a significant difference between DHH and hearing students’ scores relative to instructor-identified main points (Dowaliby & Lang, 1999). Yet, both groups of students scored only in the range of 46%–65% on that dimension, suggesting that neither group of students fully grasped what it was that the instructors would have wanted their students to learn. At the same time, the replication of significant differences on multiple-choice test scores suggests that although reading led to better performance on the part of DHH students, we have not yet determined a way to close the gap between them and hearing students. The following analyses considered whether DHH students were more aware of their learning when materials were signed or presented as text.

Metacognitive judgments. Metacognitive judgment scores were computed as in Experiment 1, and the results from each passage were analyzed using a 2 (reading status) × 2 (read vs. signed/spoken) × 2 (underestimation vs. overestimation) design in which the last factor was within subjects. For both the Mendel and the circadian rhythms material, significant main effects again were obtained for hearing status, $F(1, 36) = 13.84$, $MSE = 2.35/F(1, 36) = 7.53$, $MSE = 2.04$, and prediction accuracy, $F(1, 36) = 7.13$, $MSE = 2.15/F(1, 36) = 19.33$, $MSE = 2.25$. As in Experiment 1, hearing students were more accurate than DHH students, and overestimations of accuracy were more likely than underestimations. The Mendel material also yielded a significant main effect of condition, $F(1, 36) = 17.29$, as reading led to lesser accuracy for both DHH and hearing students. A marginal interaction of hearing status by prediction accuracy, $F(1, 36) = 3.64$, $p = .06$, reflected the fact that whereas hearing students underestimated and overestimated their performance with the Mendel material equally often, DHH students were twice as likely to overestimate as to underestimate their performance (see Table 2).

Taken together, the above results indicate that DHH students learned more but were less able to assess their learning when materials were presented as text rather than signed. This result could be a consequence of students’ having the opportunity to reread the printed material—in contrast to the transitory nature of signed and spoken language—although such a strategy is not commonly found among DHH readers (Spencer & Marschark, 2010, chapter 5). In any case, given the generally acknowledged difficulties of DHH students in reading, better learning from text than sign language is contrary to what many educators and investigators would expect (but see Marschark et al., 2009). Together with findings from Experiment 1, the results suggest that neither being provided with a summary of class content nor actively generating the main points associated with it offers DHH students sufficient tools to allow them to acquire as much of that content as their hearing peers (cf. Dowaliby & Lang, 1999). This issue was considered from a different perspective in Experiment 4. First,
however, Experiment 3 offered a somewhat more rigorous evaluation of the utility of scaffolding by using the instructors-generated main points as support for students’ reading the Mendel and circadian rhythms materials.

As another possible aid to learning, an alternative condition offered students a list of key vocabulary/definitions for each passage. Use of vocabulary support lists with DHH students has been suggested by educators like Stewart and Kluwin (2001, p. 12) and instructors at NTID and various schools for the deaf. Although this strategy is popular among both instructors and DHH students (Lang, McKee, & Conner, 1993), there does not appear to have been any empirical evaluation of its benefit to academic outcomes. Given previous findings indicating that the organization of concept knowledge among DHH students is significantly more variable and idiosyncratic than it is among hearing students (Marschark, Convertino, et al., 2004; McEvoy et al., 1999), it seems unlikely that such lists would be generally useful for groups of DHH students. Any given DHH student may be familiar with one meaning or sense of a word and unaware of its inappropriateness in any particular context, but this will vary widely even within a single DHH classroom. This is not just an issue for text but also occurs in the reception of signed materials (see Marschark et al., 2009, for examples).

Experiment 3

Method

Participants. Forty-two DHH and 42 hearing students participated, all drawn from the same pool as the previous experiments.

Materials and procedure. The two passages utilized in the previous experiment were used here, but both passages were read. Before reading one passage, students were given the list of main points generated by the instructor content experts in Experiment 2. Students were informed that they could use the list to support their comprehension and learning of the material, and they were able to view it before, during, and after reading the passage (i.e., while taking the multiple-choice test). After taking the test, each student predicted their accuracy for each question, as in the previous experiments. Before seeing the other passage, the same students were given a list of vocabulary drawn from the passage (53 for the Mendel passage and 55 for the circadian rhythms passage), each with a definition appropriate to the context. That vocabulary had been identified by students in a preliminary study in which participants were asked to circle any words in the passages that they thought “would be difficult for high school students” reading the passages. DHH students specifically were requested to think about DHH high school students.

Passage, support materials, and lecture order were balanced across small groups, tested separately with regard to hearing status. Testing was conducted by senior sign language interpreters who signed and/or spoke, as appropriate, according to student preferences.

Results and Discussion

Multiple-choice tests. Separate analyses of test scores for the two passages involved 2 (hearing status) × 2 (scaffolding materials: main points or vocabulary) analyses of variance. The Mendel and circadian rhythms passages both yielded main effects of hearing status, $F(1, 83) = 16.38, \text{MSE} = 0.02$, and $F(1, 83) = 29.15, \text{MSE} = 0.02$, respectively, in favor of the hearing students. There were no other significant effects (see Table 3). Overall, DHH student scores were approximately 82% of those obtained by hearing students.

Metacognitive judgments. Students’ predictions of their accuracy on the multiple-choice test were scored as in the previous experiments. Both the Mendel and the circadian rhythms passages yielded main effects of hearing status, $F(1, 80) = 24.94, \text{MSE} = 1.76/ F(1, 80) = 27.69, \text{MSE} = 1.98$, as hearing students were significantly more accurate in their judgments than DHH students. There were also significant main effects of prediction accuracy, $F(1, 80) = 15.34, \text{MSE} = 3.00/F(1, 80) = 56.10, \text{MSE} = 2.78$, as over-estimates of performance were greater than underestimates (see Table 3).

Comparing Experiments 2 and 3. Experiment 3 sought to close the gap between DHH and hearing students’
learning and their metacognitive judgments of learning by providing advance scaffolding support via instructor-determined main ideas and vocabulary identified by other students as potentially problematic. Overall, neither intervention appeared to be very successful. Because the materials were the same as those used in Experiment 2, however, a comparison with the (no support) reading condition of Experiment 2 offered the opportunity to evaluate the relative contributions of the support materials to both learning and metacognition using a 2 (hearing status) × 3 (support materials: main points, vocabulary, or none) between-subjects design.

Analyses of the multiple-choice scores for the separate passages yielded main effects of hearing status for both the Mendel and the circadian rhythms passages, \( F(1, 118) = 31.33, MSE = 0.22/F(1, 118) = 50.58, MSE = 0.020 \), both in favor of the hearing students (see Tables 2 and 3). Neither passage yielded a significant effect of condition, reflecting that the provision of scaffolding in the form of either instructor-generated main points or relevant vocabulary did not improve test scores.

Analyses of the metacognitive judgments scores across experiments included prediction accuracy (underestimation and overestimation) as a within-subjects factor. Analysis of both passages yielded significant main effects of hearing status, \( F(1, 118) = 3.96, MSE = 2.39/F(1, 118) = 15.08, MSE = 2.03 \), and prediction accuracy, \( F(1, 118) = 22.50, MSE = 2.79/F(1, 118) = 69.46, MSE = 2.66 \), as well as significant interactions of hearing status by support materials, \( F(2, 118) = 10.40/F(2, 118) = 5.99 \). As can be seen in Tables 2 and 3, the main effects reflected more accurate metacognitive judgments by hearing students and the finding that overestimation of performance was greater than underestimation. The interactions resulted from the fact that with both passages, hearing students were far more accurate in their judgments when their reading was supported than when it was not, whereas the DHH students were slightly more accurate when they had no support.

In summary, scaffolding the content of the two passages for DHH students not only failed to close the gap with hearing students in terms of test scores but also, as in Experiment 1, their scores with scaffolding did not reach the level of hearing students’ without scaffolding. This result suggests that differences in learning are not simply a matter of the relative availability of a framework for comprehension and that simply providing support in the form of a scaffold of the main ideas or vocabulary is not the answer. Although scaffolding did not lead to greater metacognitive accuracy for the DHH students, it did for the hearing students. That result is consistent with findings of Kruger and Dunning (1999) and Thiede and Anderson (2003), also with hearing students, indicating that greater content knowledge/comprehension is associated with greater metacognitive accuracy and, by extension, a greater likelihood of self-regulation during study. Still to be determined is why such support is effective for hearing students but not DHH students.

**Experiment 4**

Experiments 1 through 3 sought to improve DHH students’ classroom learning by providing them with

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Means and standard errors for test scores (%) and metacognitive judgments (frequencies): Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Def</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>Multiple-choice test</td>
<td></td>
</tr>
<tr>
<td>Mendel</td>
<td>Main points support 0.76 0.03</td>
</tr>
<tr>
<td>Circadian rhythms</td>
<td>Main points support 0.60 0.03</td>
</tr>
<tr>
<td>Metacognitive judgments</td>
<td>Main points support</td>
</tr>
<tr>
<td>Mendel</td>
<td>Underestimation 1.38 0.26</td>
</tr>
<tr>
<td>Vocabulary support</td>
<td>Underestimation 1.19 0.26</td>
</tr>
<tr>
<td>Circadian rhythms</td>
<td>Underestimation 1.29 0.28</td>
</tr>
<tr>
<td>Vocabulary support</td>
<td>Underestimation 1.57 0.28</td>
</tr>
</tbody>
</table>
alternative forms of scaffolding for lectures and text. This frequently employed instructional strategy may facilitate learning in some contexts, but none of the present attempts have been sufficient to bring DHH students' learning up even to the level of hearing peers who have not received such support. Nor did it significantly influence those students’ ability/tendency to accurately monitor their comprehension/learning. This is not to deny the fact that DHH students like receiving such support or that both they and their instructors believe that it is helpful (e.g., Lang et al., 1993; Stewart & Kluwin, 2001), only that there does not appear to be empirical evidence to support those assumptions. DHH students might be more likely to utilize such supports effectively, either spontaneously or with explicit instructions, if they had a stronger underpinning in the relevant content area (Rawson & Kintsch, 2002) or if they had more automatized or flexible learning strategies (Bebko, 1998).

Experiment 4 involved another attempt to facilitate DHH students' learning and metacognitive judgments through the use of scaffolding. Dunlosky and Nelson (1992) and others have shown that students’ judgments of comprehension and learning are more accurate after a delay than they are immediately following presentation of materials. Thiede and Anderson (2003), for example, examined the effects of delay on test performance and the accuracy of metacognitive judgments made by hearing students who read passages and then summarized them. They did not find an advantage in test performance as a function of delay (see also Thiede & Dunlosky, 1994), but metacognitive judgments did become more accurate. Thiede et al. (2003) conducted a similar study in which students read passages and then generated keywords associated with them (e.g., shipwreck after reading about the sinking of the Titanic). Delayed generation of keywords led to both greater judgment accuracy and higher comprehension ratings than generation immediately after reading the texts.

Thiede and Anderson (2003) explained the benefit of delay in terms of memory, leading individuals to believe that they have understood more than they really have. When judgments have to be made on the basis of long-term memory, most of the extraneous information will have decayed, and the individual will have a more accurate basis for their metacognitive judgments. In either case, the less well the material is comprehended, the greater should be the interference from extraneous information and the less accurate will be metacognitive judgments.

The Thiede and Anderson argument was founded on the relative difficulty/familiarity of the materials, but one could make a parallel argument with regard to the comprehender/learner: Students who understand less for whatever reason (e.g., language fluency, auditory or visual acuity, background knowledge) will be less accurate in their metacognitive judgments. Moore, Lin-Agler, and Zabrucky (2005) further demonstrated that metacognitive judgments made immediately after learning are strongly influenced by students’ generalized views of their abilities rather than being true assessments of their immediate performance. DHH students’ self-evaluations of their abilities—including any influence of the unskilled and unaware effect—thus could affect the accuracy of their predictions. Going a step further (although not in this article), comparison of higher achieving and lower achieving DHH students according to some external criterion (e.g., standardized achievement or college entrance tests) and perhaps even those with more of an internal versus external locus of control might yield differences in the accuracy of metacognitive judgments (Maki et al., 2005).

Experiment 4 examined the effects on metacognitive accuracy of inserting a delay between reading and subsequent testing and metacognitive judgments by DHH and hearing students. In addition, to facilitate comprehension—and hence presumably metacognitive accuracy—one of two passages read was selected so as to be highly familiar to DHH students. On the basis of the Thiede and Dunlosky arguments, both the delay and the familiarity manipulations would be expected to have greater effects on the metacognitive judgments of DHH than the hearing students. The experiment also examined the relation between key points in a passage that students deemed important.
(indicated by a highlighting task) and what they sub-
sequently generated as the main ideas. This manipu-
lation represented a modified replication of the peer
scoring method of Experiment 2 except that each
student provided their own scoring system by high-
lighting main points in each passage before being
asked to generate them.

Method

Participants. A total of 52 RIT students participated
in this experiment, 26 DHH, and 26 hearing, drawn
from the same pool as the previous experiments.

Materials and procedure. This experiment utilized
the passage about the work of Gregor Mendel used in
the previous two experiments and a passage describing
the Deaf President Now (DPN) protest by DHH staff
and students at Gallaudet University in 1988. In order
to incorporate a delay between reading and test/judgments,
the procedure followed the following sequence:
- reading the first passage,
- highlighting the four most important sentences
  in the first passage,
- reading the second passage,
- highlighting the four most important sentences
  in the second passage,
- generating up to four of the most important
  points from the first passage,
- multiple-choice test on the first passage,
- true–false test on the first passage,
- generating up to four of the most important
  points from the second passage,
- multiple-choice test on the second passage,
- true–false test on the second passage, and
- rating the familiarity of the first then the second
  passage.

The reading-to-test delay varied across students in
this self-paced sequence but did not exceed 10 min.
Instructions given during the testing session obscured
the plan for delayed testing with both passages.

Results and Discussion

A preliminary analysis compared the familiarity rat-
ings of the two passages for the two groups of partic-
ipants. As expected, the DPN passage was far more
familiar to the DHH students than the hearing stu-
dents; the Mendel passage was rated as more familiar
by the hearing students than the DHH students,
yielding a significant hearing status by passage inter-
action, $F(1, 50) = 80.46$, $MSE = 2.29$ (see Table 4).
Analyses of each set of test scores therefore included
both a 2 (hearing status) $\times$ 2 (passage: Mendel or
DPN) analysis of variance in which the second factor
was within subjects and a second analysis in which the
rated familiarity of the two passages were covariates.

Multiple-choice and true–false tests. Considering scores
on the multiple-choice test, the first analysis yielded
both a significant main effect of hearing status, $F(1,
50) = 56.42$, $MSE = 0.04$, and a significant hearing
status by passage interaction, $F(1, 50) = 24.66$, $MSE
= 0.01$. As can be seen in Table 4, although hearing
students scored higher overall, DHH students scored
higher on the DPN passage than the Mendel passage,
whereas the reverse was true for the hearing students.
The analysis of covariance still yielded a significant
main effect of hearing status, $F(1, 48) = 23.51,$

![Table 4 Means and standard errors for test scores (%), generated main ideas (%), and metacognitive judgments (frequencies): Experiment 4](https://academic.oup.com/jdsde/article/16/1/79/419086)
$MSE = 0.04$, but the significant hearing status by passage interaction was eliminated, $F(1, 74) < 1$. This finding confirms the importance of prior content knowledge on learning, particularly for DHH students. At the same time, the fact that DHH students’ scores on the more familiar (to them) DPN passage were only 79% of those obtained by the hearing students indicates that larger issues are involved.

Scores on the true–false test yielded the same results as the multiple-choice test (despite a higher probability of being correct by guessing). A significant main effect of hearing status, $F(1, 50) = 53.79$, $MSE = 0.02$, was obtained together with a hearing status by passage interaction, $F(1, 50) = 10.85$, $MSE = 0.01$, reflecting the same pattern as was obtained with the multiple-choice test. When passage familiarity was controlled, only a significant main effect of hearing status was obtained, $F(1, 48) = 17.44$. DHH students’ scores were 81% of those obtained by their hearing peers.

Main ideas. The third measure of learning consisted of the number of ideas correctly produced in the delayed generation task. Remembering that each student only had to produce four main ideas, each was scored using the three-level system used in Experiment 3. Because the two passages had different numbers of sentences (Mendel = 33 and DPN = 38), analyses involved the total scores divided by the number of sentences in the passage. Overall, students were significantly better in generating ideas from the DPN passage than the Mendel passage, $F(1, 50) = 5.37$, $MSE = 0.004$. There also was a marginal interaction of hearing status by passage, $F(1, 50) = 3.64$, $p = .06$, as DHH students performed equally well on the two passages while hearing students performed better on the DPN than the Mendel passage (see Table 4). When familiarity was controlled, the only significant effect was a hearing status by passage interaction, $F(1, 48) = 5.56$, $MSE = 0.004$. A similar analysis utilizing as the dependent variable the proportional scores of only those sentences highlighted by each participant yielded scores too low to have confidence in the analyses (all cells $<1\%$) and those scores will not be discussed further. As in Experiment 2, it appears that using student-based scoring systems does not provide any insight into performance other than demonstrating the considerable variability inherent in DHH students’ knowledge and learning strategies.

Metacognition. Analysis of students’ metacognitive judgments involved a $2 \times 2 \times 2$ (passage) analysis of variance in which the last two factors were within subjects. The first analysis yielded significant main effects of hearing status, $F(1, 50) = 53.29$, $MSE = 2.99$, and prediction accuracy, $F(1, 50) = 47.57$, $MSE = 1.90$, as hearing students were more accurate than DHH students in their judgments, and both groups were more likely to underestimate than overestimate their performance. These main effects were qualified by significant interactions of hearing status by passage, $F(1, 50) = 38.19$, hearing status by prediction accuracy, $F(1, 50) = 15.68$, and hearing status by passage by prediction accuracy, $F(1, 50) = 6.33$. As can be seen in Table 4, the interactions reflect hearing students’ being particularly accurate in their judgments on the DPN passage, in addition to being more accurate overall, whereas the DHH students were more accurate in their judgments on the DPN passage than the Mendel passage while being more likely than the hearing students to overestimate their performance. Although DHH students were more accurate with the more familiar materials, they were no more so than their hearing peers.

Comparing Experiments 3 and 4. Thiede and Anderson (2003) found that delaying students’ passage summarization resulted in more accurate metacognitive judgments although not better learning. Thiede et al. (2003) found that delaying students’ generation of keywords after reading a passage was associated with both more accurate metacognitive judgments and higher comprehension ratings, but actual comprehension/learning was not evaluated. In order to examine the effects of delay in Experiment 4, multiple-choice test scores and metacognitive judgments for the Mendel passage were compared to those obtained in Experiment 3 in which testing occurred in immediately after reading.

Multiple-choice scores were analyzed using a $2 \times 2$ (experiment) analysis of variance.
In addition to a significant main effect of hearing status, consistent with results of the separate experiments, $F(1, 135) = 96.97, MSE = 0.02$, a significant main effect of experiment was obtained, $F(1, 135) = 7.54$, as was a significant hearing status by experiment interaction, $F(1, 135) = 24.00, MSE = 1.90$. Table 5 reveals that DHH students scored significantly higher (by almost 40%) in Experiment 3, without a delay, than they did in Experiment 4, $t(66) = 4.20$. Hearing students scored significantly higher in Experiment 4, with a delay, $t(66) = 2.59$, consistent with the findings of Thiede et al. (2003), also with hearing students.

Analysis of the metacognitive judgments involved a $2 \times 2 \times 2$ (hearing status vs. experiment) analysis of variance in which the last factor was within subjects. Consistent with the findings from the analysis of multiple-choice scores, the DHH students were more accurate in their metacognitive judgments when they were made immediately, in Experiment 3, rather than after a delay. The reverse was found for the hearing students (see Table 5), again consistent with the Thiede et al. (2003) findings. This resulted in a significant hearing status by prediction accuracy interaction, $F(1, 132) = 10.86, MSE = 3.25$, in addition to main effects of hearing status, $F(1, 132) = 129.74, MSE = 1.76$, experiment, $F(1, 132) = 11.39$, and prediction accuracy, $F(1, 132) = 47.65$. Significant interactions of hearing status by prediction accuracy, $F(1, 132) = 17.16$, and experiment by prediction accuracy, $F(1, 132) = 5.04$, also were obtained. Consistent with the previously described findings, Table 5 indicates that hearing students were more accurate in their judgments than the DHH students, and overestimations were greater than underestimations, particularly in Experiment 4.

### General Discussion

The present experiments were motivated by recent findings suggesting that one barrier to greater academic achievement on the part of DHH students is a relative lack of ongoing monitoring of their comprehension and learning. A variety of studies has demonstrated that DHH students are less likely to engage in such self-evaluation during reading and thus are less likely to employ metacognitive strategies such as look-back, rereading, or identifying key sentences or ideas. The result is not only reduced understanding and assimilation of the material at hand but also a generalized tendency toward superficial processing of to-be-learned text or to-be-solved problems that can result in reduced cognitive growth.

Educators and researchers have long had a near obsession with DHH students’ reading skills (and lack thereof). Recent studies, however, have demonstrated that the challenges of vocabulary, grammar, inferencing, and relational processing observed during reading also are found in DHH students’ processing of signed and spoken language (depending on individual orientation). Such findings suggest that one reason we have made so little progress in improving DHH students’ reading over the past 50 years is that their alleged reading challenges are not really about reading (Marschark, 2009). Rather, there are cognitive differences between the two populations that affect language and learning regardless of preferred language modality and whether or not the former have cochlear implants (Hauser et al., 2008; Marschark & Watters, 2008; Pisoni, Conway, Kronenberger, Henning, & Anaya, 2010). This might be particularly important with regard to reading comprehension, a domain that

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Means and standard errors for multiple-choice test scores (%) and metacognitive judgments (frequencies) in Experiment 3 compared to Experiment 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deaf</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Mendel passage</td>
<td></td>
</tr>
<tr>
<td>Experiment 3</td>
<td></td>
</tr>
<tr>
<td>Multiple-choice test</td>
<td>0.74 0.03</td>
</tr>
<tr>
<td>Metacognitive judgments</td>
<td></td>
</tr>
<tr>
<td>Underestimation</td>
<td>1.29 0.18</td>
</tr>
<tr>
<td>Overestimation</td>
<td>2.52 0.29</td>
</tr>
<tr>
<td>Experiment 4</td>
<td></td>
</tr>
<tr>
<td>Multiple-choice test</td>
<td>0.54 0.03</td>
</tr>
<tr>
<td>Metacognitive judgments</td>
<td></td>
</tr>
<tr>
<td>Underestimation</td>
<td>1.46 0.23</td>
</tr>
<tr>
<td>Overestimation</td>
<td>5.19 0.40</td>
</tr>
</tbody>
</table>
has been found particularly sensitive to hearing loss (Goldberg & Richburg, 2004; Moeller, Tomblin, Yoshinaga-Itano, Connor, & Jerger, 2007). Taken together, such findings suggest that efforts to improve reading—and learning more broadly—need to focus more on the cognitive and metacognitive skills supporting language comprehension whether or not presented as text. Such skills largely are acquired incidentally by hearing children; for deaf children, it appears that we need to teach them more explicitly.

Across a variety of domains, including reading, verbal and nonverbal problem solving, and mathematics, a common tendency among DHH students is a tendency to focus on individual items or dimensions of a task rather than to engage in integrative or relational processing (see Hauser et al., 2008; Marschark, Convertino, et al., 2006, for reviews). During reading, for example, DHH students tend to read one word at a time, and comprehension of the larger message may be lost. Although it has not been investigated explicitly, a similar situation appears to occur during comprehension of through-the-air communication (sign or speech), related to relative language fluencies, the availability of cognitive/metacognitive resources, or both. In an effort to better understand these issues, this study sought to facilitate comprehension and learning by explicitly providing students with material support during reading and classroom lectures. The kinds of support provided are commonly utilized in DHH classrooms, but the evidence cited to support their use has come from studies involving hearing students.

The results from the present four experiments are consistent and easily summarized. First, hearing students regularly surpassed DHH students on various tests of learning including multiple choice, true–false, and generation of main ideas. Essentially, identical results were obtained regardless of whether material was read or presented in a classroom lecture. Furthermore, hearing students performed better even when the material was more familiar to the DHH students (the DPN passage in Experiment 4). Second, contrary to what many observers would expect, DHH students learned just as much or more when they read material as opposed to seeing it signed. Third, although all students tended to be more likely to overestimate than underestimate their test performance, hearing students were significantly more accurate than DHH students and, in particular, DHH students showed a greater tendency to think that they had learned more than they really had. Fourth, previous studies with hearing students suggested that greater material familiarity and a delay between study and the making of metacognitive judgments would increase prediction accuracy. The familiarity effect was obtained for both DHH and hearing students (Experiment 4), but the beneficial effect of delay was observed only for hearing students.

Most of the above findings replicate and extend earlier studies. Marschark, Sapere, et al. (2004, 2005) found that the DHH students were less accurate than hearing students in their metacognitive judgments following classroom lectures. Consistent with the previous finding that DHH students generally learn at least as much from what they read as from what they see signed (Marschark, Leigh, et al., 2006; Marschark et al., 2009), explicit examination of prediction accuracy in the present experiments revealed that they also were more accurate in judging their own learning when presented with text. The evidence thus suggests that rather than avoiding text materials with DHH students, their use may lead to better learning than sign language alone. At the same time, DHH students may be less confident with text than they are with sign language, and findings are accumulating indicating that, as with hearing students (Moore et al., 2005), perceptions of their own abilities can influence how they approach language comprehension, learning, and education.

Consistent with this suggestion, Richardson, MacLeod-Gallinger, McKee, and Long (2000) identified several differences in the approaches to learning of DHH and hearing students in the same mainstream classrooms. Richardson, Marschark, Sarchet, and Sapere (in press) found related differences in the attitudes, perceptions, and approaches of DHH students associated with being enrolled in mainstream versus separate (all DHH) classrooms at the same university. The extent to which the latter differences were a function of being in those classrooms versus resulted in
those placements is unclear. The central question, in any case, is whether we can provide students with materials or instruction that builds on their cognitive and academic strengths while accommodating related needs regardless of, but appropriate to the classroom in which they are enrolled.

The present attempts to facilitate learning and self-monitoring with scaffolding (main ideas or vocabulary lists) suggest that simply providing such material is insufficient to accomplish the above goals. Thiede et al. (2003) reached a similar conclusion in a study that involved hearing students’ generating keywords immediately after reading passages. It was only in a delayed condition, where participants had the opportunity to use the keywords to self-regulate their studying, that learning was significantly enhanced. The issue is even more central with regard to DHH students, however, because Experiment 4 and a variety of findings described earlier indicate that when given a delay, DHH students do not show the improved performance observed among hearing students. Presumably, such differences indicate that DHH students are either less aware of the potential of various cognitive strategies or less able to utilize them effectively (Marschark & Everhart, 1999; Mousley & Kelly, 1998; Richardson et al., 2000; Schirmer, 2003; Schirmer et al, 2004; Senior, 2004). So, can we teach students to generate such strategies spontaneously and hence improve learning through their own “self-regulation”?

Martin, Craft, and Sheng (2001) demonstrated that teacher-led instrumental enrichment (IE) training, twice weekly for 6 months, significantly enhanced DHH school children’s critical thinking and problem-solving skills, although they did not examine the effects of such changes on learning. Jonas and Martin (1985) had reported that IE training was associated with increased academic achievement and various cognitive measures among DHH children but did not provide enough information concerning the training program, research methodology, or the results to determine the nature of that association.

Mousley and Kelly (1998) demonstrated the potential of metacognitive training in a study intended to promote its use in mathematical problem solving by DHH college students. They used the Tower of Hanoi problem, a nonverbal task that requires multiple actions to arrange rings on a set of pegs in a prescribed order. In one experiment, they found that reading ability was related to the accuracy with which students could explain (in writing) the goals and strategies used in the task as well as their ability to solve a mathematics word problem, the answer to which followed similar logic. In a second experiment, requiring students to take at least 2 min to visualize the steps in solving the Tower of Hanoi problem led to solving it in significantly fewer moves than in a non-visualization condition. Mousley and Kelly concluded that both experiments indicated greater metacognitive involvement in problem solving after relatively limited intervention.

Fox (1994) gave DHH college students prereading exercises, study guides, and “limited summaries” of materials that they were reading in addition to explaining how each strategy should be used and how they could benefit learning. Follow-up discussions about the exercises included emphasis on the “how and why” of using metacognitive strategies during reading. Fox suggested that students learned more from texts following the semester-long training, but the only evidence presented was the (higher) grades that he assigned to students in his course. As noted earlier, his qualitative description of student performance was not encouraging.

Together with the findings from the present experiments, the above results and others like them suggest that problem solving by DHH students is likely to benefit from some kinds of metacognitive intervention, although the durability and transferability of benefits remains to be determined. With regard to reading and language comprehension at-large, the potential for training various metacognitive strategies is unclear. Educators have recommended utilizing strategies shown to be effective for enhancing reading comprehension among hearing students, but the results from the present experiments suggest that, as in other domains, manipulations that work for hearing students are not necessarily effective for DHH students. Furthermore, there do not appear to have been previous attempts to similarly influence DHH students’ learning from information acquired through sign language or spoken language. The effectiveness of either providing or encouraging students’ self-generation of the main
ideas inherent in presented materials appears to be limited at least without some kind of pretraining in their utilization. Effective use of such supporting materials requires not only insights into one’s own comprehension and learning—that is, knowing when such support is necessary—but also the spontaneous (or at least relatively rapid and effortless) linkage between what is known and what is being learned (Bebko, 1998).

To the extent that previous studies have shown that higher and lower achievers, better and poorer readers, and more and less successful problem solvers also vary in their metacognitive skills, it is clear that content knowledge and procedural knowledge must go hand-in-hand during language comprehension and learning. To the extent that DHH students generally have been found to lag behind hearing peers in both these domains, it likely that metacognitive training alone will not be optimally effective other than in content areas that are extremely familiar. The finding in Experiment 4 that although the DPN material was far more familiar to DHH than to hearing students, their test scores were still significantly lower emphasizes the content knowledge part of the equation. The failure of the Experiment 4 delay to improve their accuracy of metacognitive judgments and the general tendency among DHH students toward individual item rather than relational processing in verbal and nonverbal tasks (Marschark & Wauters, 2008) emphasizes the cognitive-procedural part of the equation.

Conclusions

In summary, there are three primary conclusions to be drawn from the present research in the context of the other recent studies that motivated it. Most importantly, findings indicating that learning (and barriers to learning) from text and sign language are much the same for DHH students offer the possibility of improving both. Emphasis on promoting metacognition and flexible comprehension strategies during reading is inherently problematic if students find reading to be an onerous and frustrating exercise. Findings to date, however, suggest that these processes might be discussed, elaborated, and enhanced first with regard to through-the-air communication and then transferred to reading.

The second general conclusion emerging from this research parallels the first but pertains to metacognition. As with learning per se, the metacognitive processes that underlie comprehension and learning from through-the-air communication are similar to those underlying learning from texts, and in both domains, DHH students appear to be less consistent or effective in self-monitoring. As with hearing students, being more familiar with the content area facilitates DHH students’ recognizing what they know and what they do not know, but simply offering the opportunity for self-regulation appears to be insufficient for significantly improving accuracy. Educators and investigators have long recognized that parents and teachers often promote instrumental dependence among deaf learners, frequently providing help and support when students would benefit more from having to discover answers for themselves. For reasons of expedience or low expectations, however, demands placed on DHH students in formal and informal educational settings are often considerably less stringent than those placed on hearing students. If we want better performance, we need higher expectations.

Finally, this study emphasizes that in the domain of metacognition, as in others, methods and strategies deemed effective for hearing students do not necessarily yield the same results with DHH students. With different and more variable language, education, and experiential backgrounds, the organization of knowledge and behavior among deaf individuals is different than it is among hearing individuals and varies more widely across individuals. The recognition and accommodation of their individual differences in no way demeans or undervalues the skills and potential of DHH students. Rather, they acknowledge the reality of growing up deaf and put educators and investigators on notice that we need to do more to understand the foundations of their learning in order to support their needs and build on their strengths.

Notes

1. “DHH students” is used generically here to refer to all participants who qualified for special support services on the basis of hearing status. More broadly, there is no evidence that hearing thresholds predict differences in academic achievement (see Convertino et al., 2009; Powers, 2003), and hearing
thresholds have failed to predict learning in dozens of experiments conducted previously by the investigators with students from the same population.

2. It could be argued that the “main ideas” generated by students represent personal judgments and cannot be scored for correctness, but such a position contradicts assumptions underlying academic assessments. DHH students’ (but not hearing students’) content retellings after reading or seeing the material signed in the Marschark et al. (2009) study revealed that some of what they “learned” was not in the presented material and in some cases was totally unrelated.

3. In all the studies of mainstream classrooms cited here, analyses of gain scores yielded the same results as postlecture test scores, and the former were not examined in this study.

4. The authors thank Richard Aslin for suggesting the student-based scoring criteria, even if it did not work.

Funding

National Science Foundation (REC-0633928 and REC-0814332).

Conflicts of Interest

No conflicts of interest were reported.

Acknowledgements

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 Sarah Remelt, Thomastine Sarchet, and Doni LaRock for their assistance. Review of this manuscript was conducted by Tane Akamatsu and two anonymous reviewers.

References


