Typing versus thinking aloud when reading: Implications for computer-based assessment and training tools

BRENTON MUÑOZ, JOSEPH P. MAGLIANO, and ROBIN SHERIDAN
Northern Illinois University, DeKalb, Illinois

and

DANIELLE S. McNAMARA
University of Memphis, Memphis, Tennessee

The goal of this study was to assess the impact of modality of production of think-aloud protocols on reading strategies. Readers in two studies spoke or typed protocols for narrative or science texts and completed comprehension tests for each text. Human judges identified the presence of paraphrasing, bridging inferences, and elaborating within the protocols. Reading comprehension skill was assessed with the Nelson–Denny test. With respect to narrative texts, paraphrasing and bridging were less frequent when readers were typing than when they were thinking aloud. With respect to science texts, less-skilled readers made bridging inferences more frequently when typing than when speaking. Conversely, skilled readers generated more paraphrases than bridges when typing thoughts but not when speaking. These results have implications for computer-based tools for reading assessment and intervention.

A growing trend in reading interventions and assessment tools is to implement them on a computer (Best, Rowe, Ozuru, & McNamara, 2005; Magliano & Millis, 2003; McNamara, Leinstein, & Boonthum, 2004; Millis, Magliano, & Todaro, 2005; Ozuru, Best, & McNamara, 2004). In this way, training can be made available to large numbers of students. Such tools are designed to foster the development or assess the occurrence of reading strategies that lead to good comprehension. As an example, iSTART is designed to foster good comprehension skills by teaching students to self-explain as they read (Graesser, McNamara, & VanLehn, in press; Magliano et al., 2005; O'Reilly, Sinclair, & McNamara, 2004). Such strategy training tools typically require students to read texts on a computer and periodically either “think aloud” or self-explain after reading specific sentences. When the students type their thoughts into the computer, their protocols are analyzed via computational algorithms that assess their quality and the presence of strategies, and the students are then provided with computer-based feedback (Best et al., 2005; McNamara et al., 2004; Millis et al., 2005; Ozuru et al., 2004).

Traditionally, however, readers have produced think-aloud or self-explanation protocols verbally (Coté &

Goldman, 1999; Ericsson & Simon, 1994; Long & Bourg, 1996; Magliano, Trabasso, & Graesser, 1999; Pressley & Afflerbach, 1995; Trabasso & Magliano, 1996a, 1996b; Whitney & Budd, 1996; Zwaan & Brown, 1996). Their thoughts are recorded, transcribed, and then analyzed by human judges who assess the comprehension strategies used by the readers. One important step in evaluating the computer-based counterparts to such methods is to assess the extent to which typing one's thoughts is comparable to producing them verbally. By *comparable*, we mean that readers typing on a keyboard produce the same comprehension strategies as they would produce when thinking aloud. One would hope that the protocols produced by typing would not be qualitatively different from those produced verbally. Although a few studies have assessed comprehension strategies when readers type their thoughts (Hausmann & Chi, 2002; Magliano et al., 2005; Magliano, Wiemer-Hastings, Millis, Muñoz, & McNamara, 2002; Millis et al., 2005), none has provided direct comparisons with what those readers would do when thinking aloud. The goal of the present study was to compare reading strategies produced by readers thinking aloud and those produced by readers typing. In our first experiment, participants thought aloud and typed their thoughts while reading narrative texts; in our second experiment, they did so while reading scientific texts.

Reading Strategies Revealed When Thinking Aloud

Think-aloud methodologies have been extensively used to assess and examine comprehension processes (Coté &

This research was funded by a Department of Education (Institute for Educational Sciences) grant awarded to J.P.M. and Keith Millis and a National Science Foundation grant awarded to D.S.M. Correspondence should be sent to J. P. Magliano, Department of Psychology, Northern Illinois University, DeKalb, IL 60115 (e-mail: jmagliano@niu.edu).

Goldman, 1999; Ericsson & Simon, 1994; Long & Bourg, 1996; Pressley & Afflerbach, 1995; Trabasso & Magliano, 1996a, 1996b; Whitney & Budd, 1996; Zwaan & Brown, 1996). With this approach, readers are instructed to report whatever thoughts come to mind as they read sentences in a text. A variation of this approach asks readers to specifically generate explanations (to self-explain) as they read (Chi & Bassok, 1989; Chi, Bassok, Lewis, Reimann, & Glaser, 1989; McNamara, 2004). Think-aloud instructions are designed to encourage participants to produce only the thoughts that are immediately available in working memory and easy to produce linguistically (Ericsson & Simon, 1994). Think-aloud protocols show how available information is used in an *effortful* search for meaning during comprehension (Bartlett, 1932; Coté, Goldman, & Saul, 1998; Graesser, Singer, & Trabasso, 1994; Stein & Trabasso, 1991); they expose conscious, strategic processing. There is also substantial evidence that these protocols reveal comprehension strategies that readers would adopt while reading silently (e.g., Magliano & Millis, 2003; Magliano et al., 1999).

Various comprehension strategies can occur in think-aloud protocols (Pressley & Afflerbach, 1995; Trabasso & Magliano, 1996a). Several taxonomies have been proposed to assess the frequency of these strategies (e.g., Coté & Goldman, 1999; Magliano et al., 2002; Trabasso & Magliano, 1996a). For example, Trabasso and Magliano (1996a) distinguished among paraphrases of the sentence just read, explanatory inferences, predictive inferences, and associative inferences. They found that in the context of narrative texts, the majority of the strategies produced by participants thinking aloud were inferences. Of these inferences, almost all were explanations of the sentence just read. In addition, Magliano et al. (1999) distinguished between text-based and knowledge-based explanations. They found that the majority of inferences were explanations based on world knowledge. Although explanations based on prior text information, or bridging inferences, occurred much less frequently than knowledge-based explanations, they were most predictive of story comprehension. These explanations provide an important basis for achieving coherence in understanding (see, e.g., Magliano & Millis, 2003; Magliano et al., 1999; Trabasso & Magliano, 1996a, 1996b).

There is some evidence that this pattern of strategies occurs for scientific texts as well (Coté & Goldman, 1999). However, Chi et al. (1989) found that paraphrases occurred more frequently than explanations for difficult physics texts, although the readers who explained more had a better understanding of the text. Paraphrasing may be a more prevalent strategy for difficult scientific texts for a number of reasons. First, it may be the case that paraphrasing helps readers better comprehend the sentence that has just been read. Second, it may be the case that readers lack the domain knowledge necessary to generate inferences and as a consequence produce paraphrases when asked to think aloud. Indeed, readers who possess more domain-relevant knowledge tend to explain more of what they read (Chi et al., 1989). A third possibility is that

a paraphrase provides a basis for beginning an explanation of a more difficult text (McNamara, 2004; Todaro, Magliano, Millis, McNamara, & Kurby, 2005).

To our knowledge, no studies have directly compared protocols produced by thinking aloud and those produced by typing. However, a few studies have had readers produce self-explanation protocols via typing that suggest that the modality of producing protocols may affect the use of reading strategies (Hausmann & Chi, 2002; Magliano et al., 2005; Magliano et al., 2002; Millis et al., 2005). Magliano et al. (2002) examined self-explanation protocols and found that approximately 65% of the clauses were paraphrases, which is more than what is typically found when readers think aloud. Even more dramatic, Hausmann and Chi (2002) found that typed self-explanation protocols were dominated by paraphrases, which contrasts with what Chi and her colleagues have reported for thoughts produced verbally (Chi et al., 1989).

There are a number of reasons why typing may change reading strategies in comparison with thinking aloud. First, the cognitive demands of typing may place a sufficient burden on working memory such that inferences become less available to the reader. Readers will thus focus on the most immediately available context—namely, the current sentence. Also, readers may be more apt to edit their thoughts when typing. In doing so, they may focus more on the current text context than on inferences.

Given the goal of the present study, an obvious shortcoming of previous studies was that they did not provide within-participants comparisons of spoken and typed protocols. Another potential problem is that not every sentence affords inferences for one who is thinking aloud. For example, Magliano et al. (1999) found that readers tended to generate causal bridging inferences (i.e., text-based explanations) when there were implied causal relationships between the target sentence and a prior text sentence or sentences. Conversely, readers tended to generate elaborative inferences based on world knowledge when there was a causal break in the story or when new entities were introduced. In the present study, we carefully selected the sentences for which readers produced verbal protocols so that they afforded inferences. That is, we chose sentences that contained implied causal relationships with a prior text sentence.

We report two experiments. In Experiment 1, participants thought aloud and typed their thoughts while reading relatively simple narrative text. In Experiment 2, they did so for more difficult scientific texts. Narrative and scientific texts place different burdens on the reader. With respect to narrative texts, readers have extensive background knowledge that they can draw upon in order to comprehend the text and when they produce think-aloud protocols. This is one reason why the majority of inferences produced by readers thinking aloud in response to narrative are knowledge-based elaborations (e.g., Trabasso & Magliano, 1996a). Conversely, readers typically do not have extensive knowledge for scientific text and are expected to produce fewer elaborative inferences than would be observed for narrative texts. It is therefore important to

have a comparison of narrative and scientific texts with respect to typing and thinking aloud one's thoughts. Also important is that we assessed differences between skilled and less skilled readers in both experiments. Given that many computer-based interventions are designed to help less skilled readers, it is important to assess how they respond to typing relative to thinking aloud.

EXPERIMENT 1

In Experiment 1, participants produced spoken and typed protocols while reading relatively simple narrative texts.

Method

Participants. Forty-nine Northern Illinois University undergraduate psychology students participated for course credit. The data from 45 were analyzed. Four participants' data were not included in the analyses because of inaudible voice recordings.

Design. The design was a skill (skilled vs. less skilled) $\times$ mode (spoken vs. typed) $\times$ strategy (paraphrase vs. bridge vs. elaboration) mixed design. Strategy and mode were within-participants variables, and reading skill was a between-participants variable. Reading skill was assessed with the Nelson-Denny Test of Reading Comprehension (Form F), a standardized test of reading comprehension that involves the reading of short texts and the answering of multiple-choice questions. A median split was used to determine skilled and less skilled readers. Participants at the median or plus or minus one point from it were excluded from the analyses. Skilled readers ($M = 29.36$, $SD = 4.45$) had a higher Nelson-Denny score than did less skilled readers ($M = 16.76$, $SD = 3.46$) [$t(38) = 5.62$, $p < .05$].

Materials. Two Chinese folktales analyzed by Trabasso and colleagues (Trabasso & Sperry, 1985; Trabasso & van den Broek, 1985) were used in this study: "How to Fool a Cat" and "The Squire's Bride." These narratives were chosen, in part, because they were relatively simple and well within the reading skill level of college students. "How to Fool a Cat" was 39 sentences long, with a Flesch-Kincaid grade level of 3.2. "The Squire's Bride" was 43 sentences long, with a Flesch-Kincaid grade level of 3.4.

Procedure. Participants were run individually in a small room containing a personal computer. They were first administered the Nelson-Denny Test of Reading Comprehension. They were given 15 min to complete the test. Next, the participants read two short narratives and produced verbal protocols as they did so. They were instructed to report their thoughts as they read two short stories. They were instructed to do so either verbally or by typing their thoughts. They were instructed that they were to report the thoughts that immediately came to mind when they comprehended the sentence that they had just read (see Trabasso & Magliano, 1996b for similar instructions). The participants were told that they would think aloud only to certain sentences. For each text, the participants were prompted to produce protocols at 12 sentences that were roughly equally distributed throughout the texts. The participants either typed or orally produced their thoughts for an entire text. The order of the modality for producing the protocol and text assignment to the two conditions were counterbalanced. Microsoft Excel was used to present the texts and to collect the verbal protocol data. The texts were presented sentence by sentence on the screen, and the participants typed their self-explanations into a box that appeared below each target sentence. The background of the screen was white, and the text was presented in black font. The participants pushed a button at the bottom of the screen that was marked "next" to advance to the next sentence, which appeared at the bottom of the screen. Access to the Excel toolbars was removed so that the participants could proceed only by pressing the "next" button. Paragraph formatting was maintained in the presentation of the text so that the text would look natural to the participants. The participants could use

the scroll bar to reread any portion of the text that was not visible on the screen. The prompts to produce a protocol were in red and read "Please report your thoughts now." As thoughts were being typed, a box appeared on the screen just above the "next" button. This button was pressed after the participants produced their thoughts, which were recorded into an Excel spreadsheet. When orally producing their thoughts, the participants were instructed to say their thoughts out loud, and their utterances were recorded by a tape recorder. The tape recorder was turned on when the participants started reading the text assigned to this condition.

During the recall phase, the participants were given two sheets of paper containing the titles of the two texts. They were instructed to write down everything that they could remember about each text. Finally, a self-report measure of typing skill was administered. The scale for the measure allowed the participants to rate their typing skill as poor, limited, good, or excellent. Very few participants (three) reported that their typing skill was less than "good."

Analysis of the spoken and typed protocols and recall protocols. The verbal protocols were scored by trained human judges using a coding system designed to identify the participants' comprehension strategies—specifically, paraphrases, bridges, and elaborations. Paraphrases involved producing information from the current sentence. There were three levels for analyzing paraphrases. A "0" indicated that no paraphrase was present. A "1" indicated that the protocol contained a noun or noun phrase from the current sentence. A "2" indicated that the protocol contained a verb clause based on the content of the current sentence. Bridges were instances where readers mentioned information from the prior text. A "0" indicated that the protocol did not contain a bridge. A "1" indicated that the protocol contained a noun or noun phrase from any prior sentence. A "2" indicated that the protocol contained a verb clause based on prior text information. Finally, elaborations were inferences that contained information not mentioned in the text. A "0" indicated that no elaboration was present. A "1" indicated that the protocol contained a noun or noun phrase not present in the text. A "2" indicated that the protocol contained a verb clause from world knowledge. It is important to note that these strategies were not mutually exclusive and that the protocols could contain any combination of these strategies. Interrater reliability for assessing the presence of each strategy was acceptable ($K = .73$ to $.78$).

In order to score the recall protocols, the texts were parsed into verb clauses. The participants' recall protocols were also parsed into verb clauses. The verb clauses in the texts were scored to determine whether they were present in the recall protocol. Interrater reliability for these judgments was acceptable ($K = .89$).

Results and Discussion

There were two sets of analyses. The first involved the strategy scores for the spoken and typed verbal protocols, and the second involved an analysis of the recall protocols. Average strategy scores for each participant, which could range from 0 to 2, and the means as a function of modality, strategy, and reading skill are presented in Table 1. A skill (skilled vs. less skilled) $\times$ mode (spoken vs. typed) $\times$ strategy (paraphrase vs. bridge vs. elaboration) mixed ANOVA was conducted. There was a main effect of reading strategy [$F(2,70) = 116.85$, $MS_e = .078$, $p < .05$]. Post hoc analyses (LSD) revealed that there was a higher strategy score for elaborations ($M = 1.57$, $SD = .81$) than for paraphrases ($M = 1.00$, $SD = .75$), which in turn had a higher score than did bridges ($M = .96$, $SD = .77$). There was also a main effect of modality of protocols such that there were higher strategy scores from spoken protocols ($M = 1.22$, $SD = .53$) than from typed protocols ($M = 1.14$, $SD = .46$) [$F(1,35) = 13.13$, $MS_e = .11$,

$p < .05$]. Most important for the purposes of the present experiment, there was a significant strategy $\times$ mode interaction (see Figure 1) [$F(2,70) = 5.66$, $MS_e = .06$, $p < .05$]. Post hoc comparisons (LSD) revealed that strategy scores for paraphrases and bridges were lower for typing than for thinking aloud. However, there was no difference in the strategy scores for elaborations as a function of modality of producing the protocols. No other effects were significant (all $ps > .05$).

A skill (skilled vs. less skilled) $\times$ mode (spoken vs. typed) $\times$ strategy (paraphrase vs. bridge vs. elaboration) mixed ANOVA was conducted on the proportion of text clauses recalled. There were no significant main effects or interactions ($p < .05$). Apparently, modality for producing protocols did not have an impact on the comprehension of narrative texts.

The results of this experiment indicate subtle changes in reading strategies as a function of modality when readers comprehend relatively simple narratives that are far below expected reading competencies (i.e., the participants were college students, whereas the texts had a third grade reading level). Paraphrases and bridges occurred to a lesser extent for typing than for speaking, but there were no differences in elaborations. Most importantly for computer-based interventions, there were no differences as a function of reading skill. Any differences as a function of reading skill would be important for these interventions, because they are typically designed to help struggling readers. It is important to note that modality did not have an impact on our measure of comprehension, recall. However, given the simple nature of these texts, one might not expect there to be an impact. In Experiment 2, we used scientific texts that were just below the grade level of the participants and that provided a better assessment of the impact of modality on comprehension.

EXPERIMENT 2

In Experiment 2, participants produced spoken or typed protocols while reading relatively more difficult scientific texts. Short-answer questions were used to assess memory for the scientific texts. These questions were designed to assess memory for explicit text content and the underlying situation model. We did not use short-answer questions for the narratives because of the difficulty of developing

Figure 1. The modality $\times$ strategy interaction for narrative texts.

questions that assessed the deeper meaning of such simple narratives.

Method

Participants. Fifty-nine Northern Illinois University undergraduate psychology students participated for course credit. Data from 5 of the participants were not included in the analyses because of inaudible voice recordings or missing data.

Design. The design was a skill (skilled vs. less skilled) $\times$ mode (spoken vs. typed) $\times$ strategy (paraphrase vs. bridge vs. elaboration) mixed design. Again, Nelson–Denny scores were used to identify skilled and less skilled readers using a median split (i.e., participants who were at or one point away from the median were removed from the analyses). Skilled readers ($M = 30.54$, $SD = 3.30$) had higher Nelson–Denny scores than did less-skilled readers ($M = 20.42$, $SD = 3.80$) [$t(48) = 9.91$, $p < .05$].

Materials. Two expository texts were used: “The Transportation of Heat” and “Plant Growth and Development.” These texts were chosen because of their relative difficulty. “The Transportation of Heat” was 35 sentences long, with a Flesch–Kincaid reading level of 10.5. “Plant Growth and Development” was 40 sentences long, with a Flesch–Kincaid reading level of 10.7.

Ten short-answer questions per text were written to assess understanding of the explicit text base and situation model (van Dijk & Kintsch, 1983). The text-based questions tapped memory for the context of single sentences. The situation model questions assessed causal inferences, content of text events, or applications of the knowledge from the text. Situation model questions assess a deeper level of comprehension than do text-based questions (McNamara, 2004; van Dijk & Kintsch, 1983). There were five questions of each type per text. The answers for text-based questions could be found within a single sentence within a text. In contrast, the answers to bridging questions required readers to draw upon information from multiple sentences, establish causal relationships between text sentences, or engage in higher order reasoning. The ideal answers for each question were identified. These answers could require multiple parts. Each part of an answer was identified. The completeness of an answer was determined by whether participants provided information for each part and was calculated by a proportion score (i.e., the number of parts provided in an answer by a participant divided by the total number of parts of that answer) for each of the answers to determine the completeness of the answer.

Procedure. There were three phases to the experiment: The Nelson–Denny test phase, the think-aloud phase, and a short-answer test

Table 1
Mean Strategy Scores and Standard Deviations for
Think-Aloud Strategies Across Presentation Mode and
Reading Skill for Narrative Texts

Skill	Mode	Strategy					
		Paraphrase		Bridge		Elaboration	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Less skilled	Spoken	1.20	.28	1.07	.36	1.56	.34
	Typed	.95	.31	.93	.49	1.58	.41
Skilled	Spoken	1.18	.31	1.02	.33	1.69	.27
	Typed	.82	.22	.88	.24	1.62	.33

phase. Phases 1 and 2 (Nelson–Denny and think aloud) followed the same procedure as in Experiment 1, though expository texts were presented in Phase 2. During the test phase, participants were given a 10-question short-answer test about each text. Text order for questions was counterbalanced. As in Experiment 1, the participants were given a self-report question regarding typing skill. Again, very few participants (4) reported that their typing skills were less than “good.”

Analysis of the protocols and short-answer questions. Protocols were scored using the same scoring system as in Experiment 1. The short-answer questions were scored on the basis of the ideal answers. The completeness of an answer was determined by whether participants provided information for each part and was calculated by a proportion score as defined above.

Results and Discussion

There were two sets of analyses. The first involved the strategy scores for the spoken and typed protocols, and the second involved the performance on the short-answer questions. The mean strategy scores as a function of modality, strategy, and skill are presented in Table 2. A skill (skilled vs. less skilled) $\times$ mode (spoken vs. typed) $\times$ strategy (paraphrase vs. bridge vs. elaboration) mixed ANOVA was conducted on the data with skill as the between-participants variable. There was a main effect of strategy [$F(2,94) = 3.54$, $MS_e = .20$, $p < .05$]. Post hoc tests (LSD) revealed that strategy scores for paraphrases ($M = 1.15$, $SD = .40$) were higher than those for bridges ($M = .99$, $SD = .43$) or elaborations ($M = 1.01$, $SD = .49$), which did not differ.

This main effect was qualified by a significant strategy $\times$ mode $\times$ skill interaction [$F(2,94) = 3.59$, $MS_e = .082$, $p < .05$]. Post hoc analyses (LSD) revealed that when speaking, less skilled readers had higher strategy scores for paraphrases than for bridges or elaborations, which did not differ. On the other hand, there were no differences in the strategies for skilled readers when speaking. A critical difference between reading skills was that skilled readers had higher strategy scores for bridges than did less skilled readers. This replicates results from a number of studies that have investigated differences in reading strategies for skilled and less skilled readers thinking aloud (e.g., Magliano & Millis, 2003). With respect to typing, there were no differences in strategy scores for less skilled readers. Indeed, there was an increase in bridging inference scores for typing relative to speaking, but this only approached significance ($p < .06$). No other differences approached significance across the modalities for less skilled readers. Unlike when thoughts were spoken, skilled readers had higher paraphrase scores than bridging scores when typing. However, there were no significant differences in the strategy scores across the different modalities for these readers. A critical finding for skilled and less skilled readers was that the strategy scores for bridges did not differ between the two groups when they were typing, as opposed to when they were speaking. Indeed, when less skilled readers typed their thoughts, their strategy scores for bridges were as high as those of skilled readers speaking their thoughts.

An analysis was conducted on accuracy on short-answer questions. Specifically, a skill (skilled vs. less-skilled) $\times$ question type (text base vs. situation model) $\times$ modality (spoken vs. type) analysis was conducted on the percentage of questions that were answered correctly. This analysis revealed a main effect of question type: A higher percentage of text-based questions ($M = .91$, $SD = .38$) than of situation model questions ($M = .75$, $SD = .37$) was answered completely [$F(1,46) = 56.89$, $MS_e = .02$, $p < .05$]. This main effect is consistent with the results of prior research in which text-based questions have been found easier to answer than situation model questions (e.g., Magliano et al., 2005). No other effects were significant (all $ps > .10$). Again there was no evidence that modality of producing the verbal protocols affected comprehension.

Differences were found in reading strategies as a function of modality for scientific texts. Furthermore, these differences were mediated by reading skill. Less skilled readers showed higher bridging scores when typing than when speaking. Conversely, skilled readers appeared to paraphrase more when speaking. As was the case with simple narratives, the modality of think-aloud did not affect comprehension measures. These results have important implications for computer-based reading skill interventions; they suggest that even for relatively difficult texts, typing thoughts while one is reading does not hinder comprehension processes.

GENERAL DISCUSSION

This study was designed to assess differences in reading strategies as a function of speaking or typing think-aloud protocols. We assessed this when readers comprehended relatively simple narrative text and more difficult scientific text. Whereas the narrative texts in Experiment 1 were well below the grade level of the participants, the science texts in Experiment 2 were more difficult and were closer in terms of the participants’ grade level. Besides the general difficulty of these different genres, there are also differences in readers’ relevant background knowledge. Readers have considerably more relevant background knowledge to support comprehension processes for reading narratives than for reading scientific texts (Graesser,

Table 2
Mean Strategy Scores and Standard Deviations for
Think-Aloud Strategies Across Presentation Mode and
Reading Skill for Science Texts

Skill	Mode	Strategy					
		Paraphrase		Bridge		Elaboration	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Less skilled	Spoken	1.16	.47	.89	.52	.98	.53
	Typed	1.14	.45	1.04	.49	.96	.46
Skilled	Spoken	1.13	.47	1.13	.56	1.07	.53
	Typed	1.18	.45	.92	.48	1.03	.45

1981; Graesser & Clark, 1985; Graesser, Golding, & Long, 1996). These differences appeared to have an effect on strategies produced when speaking or typing while comprehending narrative and science texts.

Across both experiments, the changes in reading strategies could be described as subtle, rather than dramatic as they were in Hausmann and Chi (2002). With respect to narrative texts, we found modest decreases in paraphrase and bridging scores for typing rather than thinking aloud. However, there were no changes in elaboration scores as a function of modality. Prior research has shown that elaborative inferences are most prevalent when readers think aloud in response to narrative texts (Magliano et al., 1999; Trabasso & Magliano, 1996b). It is important to note that the modality of thinking aloud did not change the occurrence of this dominant strategy. The knowledge that supports elaborative inferences may be activated automatically with little conscious effort on the part of the reader (McKoon & Ratcliff, 1992). The results of Experiment 1 show that the modality of thinking aloud does not hinder the production of elaborative inferences. Conversely, paraphrases and bridging inferences may require more effort on the part of the reader, which may explain the decrease in their production when readers type their thoughts. Another way to look at this is to say that thinking aloud is less effortful and less time consuming, and thus less "costly" for verbalizing a thought. Thus, readers may be more likely to produce paraphrases when speaking because it costs so little to do so. When readers are typing, they may be less likely to express these text-bound thoughts and instead express thoughts that go beyond the text and are less obvious. We can refer to this as an *economy of expression* in typing. This economy of expression implicitly indicates that readers know that paraphrases add less to their understanding of the text than do elaborations, and thus spend less time verbalizing those thoughts when typing than when speaking.

When reading science texts, less skilled readers seemed to benefit from typing their thoughts rather than speaking them. That is, they bridged to the same extent as did skilled readers when either speaking or typing their thoughts. This finding is important, because bridging inferences are a critical basis for establishing coherence in understanding (Graesser et al., 1994). In addition to economy of expression, typing may offer greater time for reflection. Less skilled readers seemed to have reaped the benefits of this more reflective response mode, and they were able to express more inferences when typing than when speaking. On the other hand, skilled readers bridged less often than they paraphrased when typing as opposed to speaking. Although this result is consistent with results reported by Hausmann and Chi (2002), one could not characterize these changes as an indication that typed protocols are dominated by paraphrasing as those researchers found.

It is important to note that the difficulty of the texts used should have been well within our participants' reading skill levels for both experiments. The narratives were at a 3rd grade reading level, and the scientific texts were

around a 10th grade reading level. It might be that changes in reading strategies would be more dramatic and consistent with Hausmann and Chi (2002) for texts at or above the competency of the readers.

These results have important implications for reading skill interventions (e.g., iSTART, Graesser et al., in press; Magliano et al., 2005; O'Reilly et al., 2004) and assessment tools (Magliano & Millis, 2003; Millis et al., 2005) implemented on a computer. Such tools require readers to type either thoughts or explanations while reading. An important question has been whether typing changes or hinders the reading process in comparison with thinking aloud. The present results indicate that typing does not appear to lead to dramatic changes in reading strategies or comprehension in comparison with thinking aloud. If anything, typing may permit readers' expression of inferences as opposed to text-bound processes such as paraphrases. Given that comprehension did not change as a function of modality, one might suppose that the inference processes occurred regardless of modality, but that typing allowed readers the time to access and express these thoughts. Thus, given the potential complications of adding voice recognition to computer systems eliciting think-aloud protocols, these results are highly encouraging.

A number of other factors may influence strategy production when verbal protocols are typed. Typing skill is one of them. Low-skilled typists would likely have difficulty generating reading strategies that demonstrate comprehension abilities while typing. One would not know whether the activity of typing suppressed the generation of these strategies or whether they were being omitted from the protocol because of difficulty in producing them. In the present study, a self-report measure of typing skill was collected. However, very few participants reported having typing skills less than "good." We were therefore unable to explore this important issue. The results of the self-report measure do indicate that the typing skills of college students are generally sufficient for the use of computer-based assessment and training tools that require typing. However, future research should explore this issue with a performance-based measure of typing skill, rather than self-report. Another factor is working memory capacity, which places constraints on inference processes (Just & Carpenter, 1992; Whitney, Ritchie, & Clark, 1991). One would expect that readers with low working memory capacity might have difficulty juggling the demands of comprehension and typing their thoughts. Further research is clearly needed to address these issues, for they are of critical importance for computer-based training and assessment tools that require readers to type verbal protocols.

REFERENCES

- BARTLETT, F. C. (1932). *Remembering*. Cambridge: Cambridge University Press.
- BEST, R. M., ROWE, M., OZURU, Y., & McNAMARA, D. S. (2005). Deep-level comprehension of science texts: The role of the reader and the text. *Topics in Language Disorders*, 25, 65-83.
- CHI, M. T. H., & BASSOK, M. (1989). Learning from examples via self-

- explanations. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 251-282). Hillsdale, NJ: Erlbaum.
- CHI, M. T. H., BASSOK, M., LEWIS, M. W., REIMANN, P., & GLASER, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, **13**, 145-182.
- COTÉ, N., & GOLDMAN, S. R. (1999). Building representations of informational text: Evidence from children's think-aloud protocols. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 169-193). Mahwah, NJ: Erlbaum.
- COTÉ, N., GOLDMAN, S. R., & SAUL, E. U. (1998). Students making sense of informational text: Relations between processing and representation. *Discourse Processes*, **25**, 1-53.
- ERICSSON, K. A., & SIMON, H. A. (1994). Verbal reports on thinking. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 24-53). Clevedon, U.K.: Multilingual Matters.
- GRAESSER, A. C. (1981). *Prose comprehension beyond the word*. Norwood, NJ: Ablex.
- GRAESSER, A. C., & CLARK, L. F. (1985). *Structures and procedures of implicit knowledge*. Norwood, NJ: Ablex.
- GRAESSER, A. C., GOLDING, J. M., & LONG, D. L. (1996). Narrative representation and comprehension. In R. Barr & M. L. Kamil (Eds.), *Handbook of reading research* (Vol. 2, pp. 171-205). Hillsdale, NJ: Erlbaum.
- GRAESSER, A. C., McNAMARA, D. S., & VANLEHN, K. (in press). Scaffolding deep comprehension strategies through Point&Query, Auto-Tutor, and iSTART. *Educational Psychologist*.
- GRAESSER, A. C., SINGER, M., & TRABASSO, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, **101**, 371-395.
- HAUSMANN, R. G. M., & CHI, M. T. H. (2002). Can a computer interface support self-explaining? *Cognitive Technology*, **7**, 4-14.
- JUST, M. A., & CARPENTER, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, **99**, 122-149.
- LONG, D. L., & BOURG, T. (1996). Thinking-aloud: Telling a story about a story. *Discourse Processes*, **21**, 329-339.
- MAGLIANO, J. P., & MILLIS, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition & Instruction*, **3**, 251-283.
- MAGLIANO, J. P., TODARO, S., MILLIS, K. K., WIEMER-HASTINGS, K., KIM, H. J., & McNAMARA, D. S. (2005). Changes in reading strategies as a function of reading training: A comparison of live and computerized training. *Journal of Educational Computing Research*, **32**, 185-208.
- MAGLIANO, J. P., TRABASSO, T., & GRAESSER, A. C. (1999). Strategic processes during comprehension. *Journal of Educational Psychology*, **91**, 615-629.
- MAGLIANO, J. P., WIEMER-HASTINGS, K., MILLIS, K. K., MUÑOZ, B. D., & McNAMARA, D. S. (2002). Using latent semantic analysis to assess reader strategies. *Behavior Research Methods, Instruments, & Computers*, **34**, 181-188.
- McKON, G., & RATCLIFF, R. (1992). Inference during reading. *Psychological Review*, **99**, 440-466.
- McNAMARA, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, **38**, 1-30.
- McNAMARA, D. S., LEVINSTEIN, I. B., & BOONTHUM, C. (2004). iSTART: Interactive strategy training for active reading and thinking. *Behavior Research Methods, Instruments, & Computers*, **36**, 222-233.
- MILLIS, K. K., MAGLIANO, J. P., & TODARO, S. (2005). *Measuring discourse-level processes with verbal protocols and latent semantic analysis*. Manuscript submitted for publication.
- O'REILLY, T. P., SINCLAIR, G. P., & McNAMARA, D. S. (2004). iSTART: A Web-based reading strategy intervention that improves students' science comprehension. In K. Kinshuk, D. G. Sampson, & P. Isaías (Eds.), *Proceedings of the IADIS International Conference on Cognition and Exploratory Learning in the Digital Age* (pp. 173-180). Lisbon: IADIS Press.
- OZURU, Y., BEST, R., & McNAMARA, D. S. (2004). Contribution of reading skill to learning from expository texts. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (pp. 1071-1076). Mahwah, NJ: Erlbaum.
- PRESSLEY, M., & AFFLERBACH, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Erlbaum.
- STEIN, N. L., & TRABASSO, T. (1991). Children's understanding of changing emotional states. In C. Saarni & P. Harris (Eds.), *Children's understanding of emotion* (pp. 50-77). New York: Cambridge University Press.
- TODARO, S., MAGLIANO, J. P., MILLIS, K. K., McNAMARA, D. S., & KURBY, C. (2005). *Understanding factors that influence the content and form of think-aloud protocols: The roles of the reader and the text*. Manuscript submitted for publication.
- TRABASSO, T., & MAGLIANO, J. P. (1996a). Conscious understanding during comprehension. *Discourse Processes*, **21**, 255-287.
- TRABASSO, T., & MAGLIANO, J. P. (1996b). How do children understand what they read and what can we do to help them? In M. Graves, P. van den Broek, & B. Taylor (Eds.), *The first R: A right of all children* (pp. 160-188). New York: Columbia University Press.
- TRABASSO, T., & SPERRY, L. L. (1985). Causal relatedness and importance of story events. *Journal of Memory & Language*, **24**, 595-611.
- TRABASSO, T., & VAN DEN BROEK, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory & Language*, **24**, 612-630.
- VAN DIJK, T. A., & KINTSCH, W. (1983). *Strategies in discourse comprehension*. New York: Academic Press.
- WHITNEY, P., & BUDD, D. (1996). Think-aloud protocols and the study of comprehension. *Discourse Processes*, **21**, 341-351.
- WHITNEY, P., RITCHIE, B. G., & CLARK, M. B. (1991). Working memory capacity and the use of elaborative inferences. *Discourse Processes*, **14**, 133-145.
- ZWAAN, R. A., & BROWN, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes*, **21**, 289-327.

(Manuscript received November 8, 2005;
revision accepted for publication January 21, 2006.)