J. EDUCATIONAL COMPUTING RESEARCH, Vol. 32(2) 185-208, 2005

# CHANGES IN READING STRATEGIES AS A FUNCTION OF READING TRAINING: A COMPARISON OF LIVE AND COMPUTERIZED TRAINING

**JOSEPH P. MAGLIANO**
**STACEY TODARO**
**KEITH MILLIS**
**KATJA WIEMER-HASTINGS**
*Northern Illinois University*

**H. JOYCE KIM**
*Rhodes College*

**DANIELLE S. McNAMARA**
*The University of Memphis*

**ABSTRACT**

The purpose of this study was to compare the relative effectiveness of live (SERT) and computer-based (iSTART) reading strategy training. Prior to and after training, participants read scientific texts and self-explained after each sentence. They also answered comprehension questions. Students showed improvement in the quality of their self-explanations and in the performance on the comprehension questions as a function of both live and computer-based training. However, there were some differences in response to iSTART training as a function of reading skill. Specifically, less skilled readers improved their performance on text-based questions, but not bridging questions. The opposite was found for skilled readers. These results indicate that computer-based, reading-skills training is effective, but different readers may improve at different levels of comprehension.

Deep comprehension arises from an effortful attempt to construct a coherent understanding of what a text is about (Bartlett, 1932; Graesser, Singer, & Trabasso,



© 2005, Baywood Publishing Co., Inc.

1994). Skilled readers engage in specific comprehension strategies when reading difficult texts, such as explaining, using logic, and elaborating (Chi, De Leeuw, Chiu, & LaVancher, 1994; McNamara, 2004b; McNamara, Best, & Castellano, 2004). Because many students do not use these strategies (Garner, 1990) and have difficulty comprehending what they read (Snow, 2002), McNamara developed Self-Explanation Reading Training (SERT) to teach readers to use active reading strategies (McNamara, 2004a, 2004b; McNamara & Scott, 1999; O'Reilly, Best, & McNamara, 2004; O'Reilly, Sinclair, & McNamara, 2004). Self-explanation refers to the act of explaining difficult text to oneself. SERT emphasizes several strategies that improve the process of self-explanation. The strategies include using logic or world knowledge to elaborate the current sentence, making conceptual bridges among ideas in the text, and predicting what will come next in the text. Not only has SERT been shown to promote general reading comprehension, it has been shown to improve overall class performance, particularly for poor students (McNamara, 2004a; McNamara, Best, & Castellano, 2004). However, implementing SERT requires special training and teachers. Furthermore, teachers may not be willing to devote class time to non-content instruction, such as comprehension training. As such, McNamara and her colleagues have developed iSTART (Interactive Strategy Training for Active Reading and Thinking; McNamara, Levinstein, & Boonthum, 2004), which is an automated version of SERT that is delivered on the computer.

The primary purpose of this study was to compare the relative effectiveness of live SERT implemented in a classroom setting to the automated training that was implemented in the context of iSTART. More specifically, we were interested in assessing whether readers respond to computer-based training in a similar fashion to live training. It is well documented that students benefit from reading skills training presented in a live format, but there is relatively little research demonstrating that students will benefit equivalently to computer-based training. Designers of computer-based training systems must demonstrate that automated training works at least as well as live training. Furthermore, it is important to explore possible advantages of computer-based training. As such, another goal of this study was to assess how readers of varying skill respond to computer-based training. A long-term goal for iSTART is to tailor training to the needs of the individual student, which is difficult to do in the context of live SERT because it occurs in a relatively large classroom setting. We cannot adequately address this goal until we have established a body of research that determines how different kinds of readers respond to training. In this study, we assess their responses to training in terms of the changes in the strategies they use while reading and changes in performance on the comprehension measures.

In Experiment 1, SERT was incorporated into the curriculum of a college-level course on critical thinking. Participants self-explained two science passages prior to and after participating in SERT. In Experiment 2, college students volunteered to participate in SERT training in the context of iSTART, and as was the case in Experiment 1, participants self-explained prior to and after training. In both experiments, reading skill was determined via a standardized test of reading skill (i.e., The Nelson-Denny Reading Skills test). The effectiveness of training in both experiments was assessed by determining changes in the students' reading strategies that occurred as a function of training and via comprehension measures.

## SERT AND iSTART TRAINING

SERT was inspired by previous research showing the benefits of strategy instruction (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi et al., 1994; Magliano, Trabasso, & Graesser, 1999; Palincsar & Brown, 1984; Yuill & Oakhill, 1988). There are three phases to SERT training: introduction, demonstration, and practice. In the current study, these phases occurred over two 50-minute class periods. The first phase, strategy introduction, includes definitions and examples of reading strategies associated with self-explanation. Table 1 contains a list of the strategies, definitions, and examples emphasized in SERT.

The second phase is the strategy demonstration. This phase is based on the assumption that students will develop a better understanding of the reading strategies and how to implement them by viewing a model (e..g, Craig, Driscoll, & Gholson, 2004; Craig, Gholson, & Driscoll, 2002). Students watch a videotape of a student overtly self explaining after every sentence in a text. The instructor periodically stops the videotape and asks the students in the classroom to identify specific strategies employed by the model. The instructor then invites a class discussion regarding the function of the different strategies in helping the model understand the text. In the current study, phases one and two occurred during one class period.

In the final practice phase of SERT, the students work in pairs to practice the SERT strategies. Students take turns reading a text out loud as they share self-explanations. Instructors are present to assist and monitor the students. This stage is based on a growing body of research that demonstrates the effectiveness of peer tutoring (e.g., Palincsar & Brown, 1984). However, the effectiveness of this stage will be influenced by the quality of the feedback provided by the members of the dyads, which can never be completely assured.

SERT has been shown to be effective in promoting better text comprehension and overall course performance, but these effects are mediated by the degree of prior knowledge for the content area (McNamara, 2004b; McNamara & Scott, 1999). McNamara (2004b) examined the effectiveness of SERT with college students who varied in prior knowledge of science. During training, half of the participants were given SERT and practiced self-explanation with four science texts and half simply read aloud the texts. After the training phase, the two groups' ability to self-explain was compared,. They also answered text-based and bridging-inference questions about the text that they had self-explained.

Table 1. Strategies Emphasized in SERT and iSTART

| SERT strategies | Definitions | Examples |
| --- | --- | --- |
| 1. Comprehension monitoring | Being consciously aware of how well you understand the words and ideas in a sentence. | **Title of Text:** Compounds and Molecules<br><br>**Monitoring:** OK, I know that a compound means things that are mixed together, but I'm not sure what a molecule is. I'll have to watch for that information. |
| 2. Paraphrasing | Restating a sentence in your own words. | **Sentence:** A hurrican hitting land is a victim of multiple processes.<br><br>**Paraphrase:** So this is saying that a hurricane goes onto land and is affected by a number of things. |
| 3. Elaborative explanation | Using domain knowledge to explain what something is or why something is happening in the text. | **Sentence:** To collect sunlight, most leaves have large, thin, flattened sections called blades.<br><br>**Elaboration:** I know plants need sunlight to live so they collect it through their leaves. |
| 4. Logic and common sense | Using reasoning skills to explain what something is or to better understand the text. | **Sentence:** Most compounds are made of molecules.<br><br>**Logic:** . . . compounds are made of molecules. Therefore, molecules must be smaller than compounds. |

Low-knowledge readers who received SERT out performed low-knowledge controls on comprehension questions, but only on text-based questions. Furthermore, low-knowledge participants in the SERT condition also demonstrated more accurate paraphrases and inferences based on domain general knowledge than did low-knowledge controls. Perhaps most encouraging is that performance on both comprehension and protocol measures were comparable to high-knowledge readers in these respects. Similarly, McNamara (2004a) discusses results showing that the course performance of low-knowledge participants who went through SERT matched that of high-knowledge participants, whereas low-knowledge controls did not (see also, McNamara et al., 2004).

McNamara and her colleagues designed iSTART (McNamara et al., 2004) to automate the core components of SERT in a web-based environment. Specifically, iSTART contains the three phases of SERT: 1) an *introductory* explanation which presents definitions and examples of the strategies, 2) a *demonstration* of the techniques in action, and 3) an opportunity to *practice* the techniques under guidance. As with SERT, these phases occur over a relatively brief period of time. In the case of the current study, a student can potentially participate in all three phases in a single session that last a little over an hour. This is different from the classroom version, which takes place over 50 minute class periods. One reason why a computerized version is substantially shorter is that peer dyads were given an entire class period for practice. The content and feedback of practice on the computer was considerably more controlled than the live version and could be implemented in a shorter period of time.

iSTART utilizes animated agents to implement these components, which is an approach that is consistent with other intelligent tutor systems, such as Auto-Tutor (e.g., Graesser, Wiemer-Hastings, Wiemer-Hastings, Harter, & Person, 1999). Auto-Tutor employs a "talking head" that asks students questions and responds in a relatively naturalistic fashion to answers typed in a computer. The logic of using such agents is that students will become increasingly engaged to the extent that they can view the system as supporting a simulated social interaction. iSTART uses multiple full-bodied characters that carry on conversations with one another, lecture, asks and answer questions, and make requests of the user to respond to activities provided in the training environment. In so doing, these agents follow a script that was inspired by SERT. With respect to strategy introduction, there is an agent who plays the role of instructor and two agents that play a role of active learners. The introduction module is designed to promote vicarious learning and to engage the student via an interactive, question-answering exchange (McNamara et al., 2004). Specifically, these agents perform a script that simulates a classroom discussion between the instructor agent and the student agents. For example, when the instructor agent questions the student agents about strategies, they banter among themselves regarding potential answers. Student agents actively ask the instructor agents questions in order to expand upon key concepts, which may not occur during a live SERT session. Perhaps most importantly, questions and answers (by both instructor and student agents) have been carefully designed to remediate understanding and correct potential misconceptions about the reading strategies.

In addition to viewing this dialog exchange, the trainee is given brief quizzes at several points during the introduction and is advised to continue to the next section or to review the last section depending on the results. Ultimately, the introduction session is designed to invite the student to be an *active* participant in the session. It is important to note that students have control over the pace of the session and can review any component of the introduction session at any time. This is another major advantage of iSTART over live SERT.

The demonstration module of iSTART also employs multiple agents who invite the trainee to take on an active role during learning. The same basic approach is taken in iSTART as is taken in SERT. The text is read and explained sentence-by-sentence by a student character (Genie) under the supervision of a teacher character (Merlin). After Genie produces a satisfactory explanation, the teacher asks the trainee to decide which strategies were used in Genie's explanation. During live SERT, this happens only intermittently, but during iSTART this happens after every sentence. Furthermore, the trainee must be engaged during this phase, whereas it is possible for an individual student to become disengaged during live SERT. At times Merlin will ask follow-up questions that focus the trainee's attention on the explanation and the target sentence.

The demonstration module dynamically adapts to the success of the trainee by varying the style of questioning that Merlin uses. The demonstration module starts out with multiple-choice questions in which the trainee is asked to identify a strategy used in example self-explanations. If they get these questions correct, trainees may be further asked to identify the segments of a self-explanation that contains a strategy. Trainees may also be asked to identify additional strategies that might be contained in a self-explanation, which emphasizes the fact that effective self-explanations often contain multiple strategies. If trainees answer several questions incorrectly, feedback is altered such that the trainee is told that a particular strategy exists in the self-explanation and to locate the segment that contains it.

During the practice phase, trainees read a brief scientific text and are asked to self-explain after every sentence by typing their explanations into a box on the computer screen. Merlin, the teacher agent during the demonstration phase, provides feedback regarding the quality of a student's self-explanations. The interaction between Merlin and the trainee follows the pattern that the trainee has observed in the demonstration section.

There are two components for evaluating the self-explanations produced by trainees. First, a series of algorithms assess three aspects of the self-explanations that reflect their quality. Second, trainees are asked to identify the strategies that they used, in much the same manner as trainees did with the self-explanations produced by the Genie agent during the demonstration phase.

Analyzing the quality of self-explanations represents a significant computational challenge for the designers of iSTART. The first stage of an evaluation process involves determining if a self-explanation is (a) sufficiently long enough, (b) significantly different from, and (b) relevant to the topic of the sentence. It is beyond the scope of the present article to provide the details of the algorithms used to analyze the protocols, which can be found in McNamara et al. (2004; see also McNamara, Boonthum, Levinstein, & Millis, in press). However, we will summarize the essence of each of these three components of the evaluation process. With respect to sufficiency in length, the designers determine a minimal number of words appropriate for a given sentence based on the number of words in that sentence and an empirically derived number of words based on a corpus of text self-explanations. Similarity to the current sentence is based on word and word stem matching. Specifically, the content words in a self-explanation are matched against the content words in a sentence, along with close paraphrases. If a significant proportion of the words in a self-explanation match the content words associated with a sentence, then the self-explanation was deemed to be too similar. An association list of words was constructed for each sentence and used to evaluate the relevancy of a self-explanation. The association list consisted of words not in a sentence, but closely associated with it and was empirically derived from self-explanations produced by a separate group of participants. Soundex (e.g., Knuth, 1998) was used to handle misspelling for both the similarity and relevancy assessments. If a self-explanation was too short, too closely associated with the sentence, or irrelevant to a sentence, then the trainee may be prompted to modify their self-explanation. If a trainee passed all three of these tests, then the trainee may be asked to identify the strategies that they used to explain a sentence.

McNamara et al. (2004) demonstrated that the evaluation algorithms used in iSTART were sufficiently accurate in judging the overall quality of the self-explanations produced by trainees. Specifically, they had human judges and the evaluation algorithm assess the overall quality of a corpus of self-explanations on a four-point scale. The results encouragingly showed that iSTART rarely gives inappropriate feedback to the trainee. For example, those explanations categorized by human judges as relevant elaborations were misclassified as level 1 (poor, but sufficient) explanations only eight times, or in 1% of the self-explanations. McNamara and her colleagues have also explored augmenting this system with Latent Semantic analysis and have shown significant improvements in the accuracy of self-explanation assessment (McNamara et al., in press) but this hybrid system was not used in the current system.

## GENERAL READING STRATEGIES

The primary goal of this study was to evaluate changes in reading strategies that occur in readers of different skill levels as a function of training. These changes should lead to improvements in comprehension, which should in turn lead to improvements in performance on outcome measures. As described above, SERT and iSTART emphasize several strategies that are associated with self-explanation (e.g., paraphrases, logic, elaboration, prediction). Any given self-explanation may contain statements that reflect multiple strategies (e.g., Best, Ozuru, & McNamara, 2004; McNamara, 2004b; Magliano et al., 1999; Ozuru, Best, & McNamara, 2004; Trabasso & Magliano, 1996).

However, for the purposes of this study, we examine changes in what we refer to as general reading strategies (Coté & Goldman, 1999; Coté, Goldman, & Saul, 1998). General reading strategies refer to the general approach a reader may take to understand a sentence in the context of the larger text context. We distinguish between three strategies, which are reflected in example protocols contained in Table 2. Readers adopt a *global-processing strategy* when they infer how a sentence is related to the larger context and/or related to domain specific and/or relevant knowledge (Graesser et al., 1994; McNamara, 2004b). The first self explanation in Table 2 represents an example of the global-processing strategy. Readers who adopt this strategy construct maximally coherent situation models by building upon their knowledge base (Coté & Goldman, 1999; Coté et al., 1998). Readers may also adopt a *local-processing* strategy, which is characterized in the second example in Table 2. Readers adopt this strategy when they infer how a sentence is related to the immediate prior sentence or produce a minor elaboration of the sentence based on general knowledge of the world. When readers tend to produce inferences based on this strategy, they will generally construct a minimally coherent textbase representation, but will not construct a detailed situation model that reflects its overarching meaning (van Dijk & Kintsch,

Table 2. Example of Global-Processing, Local-Processing, and Sentence-Focused Strategies for the Sentences "The blood cannot get rid of carbon dioxide through the lungs. It becomes purplish, and the baby's skin looks blue." from the text *Heart Disease*

| Strategy | Self-explanation |
|---|---|
| Global-processing | This gives the impression of someone choking. When someone chokes they start to turn colors and the infant is essentially choking from the inside. The skin turning blue might have something to do with not receiving enough oxygen connected to the heart problems. |
| Local-processing | When the carbon dioxide does not escape the body the baby's skin looks blue. |
| Sentence-focused | The blood turns to a purplish color and the baby's skin turns blue. |

1983; Zwaan & Radvansky, 1998). Finally, readers may adopt a strategy that is *sentence-focused* in which they attempt to flesh out the meaning of a sentence independent of the prior text context, which is exemplified in the third self-explanation in Table 2. Readers who adopt this strategy only paraphrase the sentence, without expanding on its meaning. Readers who adopt this strategy for a majority of the sentences in a text tend to construct a memory representation that is fragmented and not coherent.

There are individual differences in the extent to which readers engage in these strategies (Best et al., 2004; Chi et al., 1994; Coté & Goldman, 1999; Coté et al., 1998; Magliano & Millis, 2003; Magliano et al., 1999; McNamara, 2004b; Ozuru et al., 2004). For example, McNamara demonstrated hat readers who have less knowledge about the text topic engage in more sentence-focused and local processes, whereas readers with greater knowledge engage in more global processing. Magliano and Millis (2003) demonstrated that readers who score low on standardized tests of reading comprehension tend to engage in more sentence focused processing, whereas skilled readers engage in more local and global processing (see also, Best et al., 2004; Ozuru et al., 2004). These differences have implications for performance on outcome measures. Readers who engage in more sentence--focused processing tend to show poorer recall and ability to answer comprehension questions that tap deep understanding (Chi et al., 1994; Coté & Goldman, 1999; Coté et al., 1998; Magliano et al., 1999; Magliano & Millis, 2003).

In the present study, we examined the influence of live SERT (Experiment 1) and iSTART (Experiment 2) on the reading strategies used by skilled and less-skilled readers as well as their comprehension for texts. Given that McNamara and colleagues have demonstrated individual differences in the response to training as a function of prior domain knowledge and knowledge of reading strategies (Best et al., 2004; McNamara, 2004b; McNamara & Scott, 1999; O'Reilly et al., 2004; Ozuru et al., 2004), we expect that there will also be differences as a function of reading skill (see also, Best et al., 2004; Ozuru et al., 2004). Specifically, low-knowledge readers benefit in constructing a more complete and accurate textbase representation, whereas high-knowledge readers benefit in constructing a more accurate situation model.

With respect to the current study, we had participants self explain prior to and after receiving SERT (Experiment 1) and iSTART (Experiment 2). Students receiving training should become more active readers and produce more relevant self-explanations after training than before. An important test of iSTART is to demonstrate that readers modify their strategies as a function of training, as is the case with live SERT. However, based on prior research on individual difference in response to training, it is quite possible that we would find skilled and less-skilled readers practicing different strategies after training. Specifically, it is reasonable to expect that skilled readers would emphasize global strategies to practice building more accurate situation models. In contrast, less-skilled readers may practice paraphrasing and local processing because these strategies help construct more accurate situation models. These expectations correspond to research sited above regarding live SERT, which demonstrates that low-knowledge readers show benefits in a more accurate textbase post training, whereas high-knowledge readers show more accurate situation models after training. We would expect similar patterns with iSTART if it functions similarly to live training.

## EXPERIMENT 1

### Methods

#### *Participants*

Twenty-nine students who were enrolled in a critical thinking course taught at Northern Illinois University participated in pre-training and post-training sessions. Participants received extra credit for participating in the experimental sessions.

#### *Measure of Reading Skill*

Students took the Nelson-Denny Reading Skills test (Brown, Fishco, & Hanna, 1993), which is a standardized test of reading skill that is appropriate for college students. Only the comprehension portion of the test was used, which takes 20 minutes to administer. The test was administered after SERT was delivered and was done so at the beginning of a "class discussion day." After taking the test, students were told that the topic of the class discussion day was the validity of standardized test to measure student abilities, but no indication of this topic was provided before students took the test. With respect to the students who participated in the pre- and post-SERT experimental sessions, the mean score on the Nelson-Denny test for the skilled readers was 31.33 out of a maximum of 38. A median split was performed (median = 32) to identify skilled and less-skilled readers and participants who scored at the median were excluded from analyses. The mean score for skilled readers was 35.19 and mean score for the less-skilled readers was 26.53. Performance on the Nelson-Denny test was statistically different between the two groups, $t(26) = 7.02$, $p < .05$.

#### *Task and Design*

The current experiment employed a Reading Skill (Skilled vs Less Skilled) by SERT (Pre vs Post) mixed design. Reading skill was a between-participants variable and SERT was a within-participants variable. With respect to the experimental task, participants self-explained after every sentence while reading two scientific texts. There were two sessions in which they performed this task. The first session occurred approximately a week before SERT was administered in the course and the second session occurred within two weeks after SERT was administered.

### *Materials*

Four scientific texts were used in the pre-SERT and post-SERT sessions. The topics of the text were heart disease, the development of coal, the development of thunderstorms, and the food chain/cycle. These texts were adapted from textbooks, popular science books, or online encyclopedias. The passages were moderately difficult to read, suitable for freshman college students, and ranged between 20 and 34 sentences long (total $N = 98$). The four texts and their order of occurrence in a session were counterbalanced across participants.

Eight true/false questions were constructed for each passage. Four text-based questions of the explicit text content and four bridging questions tapped an understanding of the situation model. The text-based questions involved verifying the veracity of a statement that could be answered by reading one sentence in a text. The bridging questions required readers to consider causal relationships between one or more sentences in a text that were implied in the texts. The correct answer for four of the questions was "true" and for four of the questions it was false.

### *Coding for Self-Explanations*

Two independent raters categorized the self-explanations as blank/vague, sentence-focused, local-processing, or global-processing. A blank/vague protocol either contained no response, was a semantically depleted utterance (e.g., "OK," "uh huh," "Yep," etc.), or an uninformative metacognitive statement (e.g., "I didn't know that," "I knew that," etc.). Sentence-focused explanations contained only clauses that were either partial or entire paraphrases of the current sentence. Local-processing explanations usually contained paraphrases as well, but included at least one word/clause that indicated either a "bridge" from the current sentence to the immediately prior sentence, or a minor elaboration based on world knowledge. Global processing explanations usually contained multiple clauses that indicated a bridge(s) to the distal text context (i.e., at least two sentences back from the current sentence) and/or major elaborations based on domain specific or domain relevant word knowledge. Furthermore, if a self-explanation contained more than one clause that reflected a SERT strategy other than a paraphrase (e.g., logic, bridge, elaboration, prediction), it was also judged to be Global-processing. Inter-rater reliability in determining reading strategies was high (Kappa = .81). The judge was blind to whether the protocols were from a pre-SERT or post-SERT session.

### *SERT*

SERT was administered in the context of a critical thinking course taught at NIU. The critical thinking course teachers students to use knowledge obtained from research in cognitive psychology to improve memory, language comprehension, argumentation, reasoning, and problem solving skills. SERT was used in the course module on improving comprehension for text. This module was one of the first topics covered in the course, and occurred after the third week of the 15-week course, and as such, is an early lecture topic.

SERT took place over two, 50-minute class sessions. The instructor followed the SERT script that was developed by McNamara (McNamara, 2004b; McNamara & Scott, 1999). This script specifies the verbatim content of the sessions and when to solicit student feedback. The first class period contained both the strategy introduction and demonstration sessions. The strategy introduction session took place over approximately the first 25 minutes of the class period. During this module, each strategy is introduced and a definition is provided (see Table 1). One to three examples were provided after each definition, which consisted of a sentence(s) from a text along with a self-explanation that was produced by a student who received and was practicing SERT. Students were provided with and allowed to keep a handout that contained the definitions and examples. The instructor encouraged the students to identify the portion of an example self-explanation that contained the strategy.

Strategy demonstration followed the instruction module and took place over approximately 20 minutes. A videotape was shown of a student who was reading a text out loud and overtly practicing the SERT strategies. Students were provided with a transcript containing the texts and the model student's self-explanations. The instructor stopped the tape after predetermined locations and solicited discussion from the students with respect to the strategies used in the target self-explanation.

Strategy practice occurred during the following class period and took up the entire period. Students were then grouped into peer dyads. They were provided with a difficult scientific text (i.e., on the topic of mycorrhizae) that contained four paragraphs. A space was provided after each sentence. Students were instructed to take turns practicing the SERT strategies. Specifically, one student in a dyad was instructed to self-explain after each sentence in an entire paragraph and the other student was instructed to observe and comment on the other student's explanation. Students alternated roles after each paragraph in the text. As such, each student in a dyad self-explained two paragraphs and observed two paragraphs. Students produced their self-explanations in writing in the space provided after each sentence. After producing a self-explanation, the partner that observed was instructed to identify the SERT strategies that (s)he believed that the partner used and the dyad discussed the self-explanation.

### *Procedure*

Participants were asked to participate in a study associated with the course during the first week of class. Participants produced self-explanations while reading two scientific texts prior to and after receiving SERT. The format for the pre-SERT and post-SERT sessions was virtually identical. The participants were instructed to type in a self-explanation after reading each sentence of two texts. 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 the current sentence. The background of the screen was white and the text was presented in black font. Participants pushed a button 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 participants could only proceed by pressing the “next” button. This button was pressed after the participants produced their self-explanations, which were recorded into an Excel spreadsheet. Paragraph formatting was maintained in the presentation of the text so that the text would look natural to the participants. Participants could use the scroll bar to reread any portion of the text that was not visible on the screen. Participants were provided with a five-sentence text in order to practice reading and typing in their self-explanations. After reading both texts in a session, participants answered true/false questions for each text, which were presented in the order of presentation of the texts. The pre-SERT and post-SERT instructions differed in one respect. During the pre-SERT sessions, participants were told to self-explain by producing whatever thoughts immediately come to mind regarding their understanding of a sentence in the context of the text. During post-SERT sessions, participants were explicitly told to practice the SERT reading strategies when producing their self-explanations. Furthermore, they were provided with a sheet that listed the strategies with their definitions to remind them of the strategies.

## RESULTS AND DISCUSSION

Two sets of analyses were conducted. One set involved the self-explanations and the second involved performance on the true/false comprehension test.

### Analysis of the Self-Explanations

In order to assess changes in strategy as a function of SERT, mean strategy scores were calculated for the pre-SERT and post-SERT self-explanations, which are shown in Table 3. A reading skill by SERT mixed ANOVA was conducted. There was a significant reading skill by SERT interaction, $F$ (1,26) = 8.12, Mse = .04253, $p < .05$. Post-hoc analyses (LSD) revealed that skilled readers had higher strategy scores after training than before. In contrast, less-skilled readers had lower strategy scores than before training, although this difference was marginal ($p < .10$). Further, prior to training, skilled and less-skilled readers had similar strategy scores, whereas after training skilled readers had higher strategy scores than less-skilled readers.

We conducted another analysis to further explore how readers of differing skill responded to training. Magliano and Millis (Magliano & Millis, 2003; Magliano,

Table 3. The Mean Strategy Score as a Function of Reading Skill and SERT

| | | SERT | |
|---|---|---|---|
| Strategy | Skill | Pre | Post |
| Experiment 1 (SERT) | Skilled | 2.18 (0.38) | 2.28 (0.32) |
| | Less-skilled | 2.11 (0.52) | 2.03 (0.40) |
| Experiment 2 (iSTART) | Skilled | 1.85 (0.44) | 2.02 (0.27) |
| | Less-skilled | 1.71 (0.45) | 1.81 (0.37) |

**Note:** Standard deviations are reported in the parentheses.

Wiemer-Hastings, Millis, Munoz, & McNamara, 2002) have demonstrated that sentence-focused strategies tend to have greater semantic overlap with the current sentence than with the prior context, whereas global strategies tend to be semantically related to both the local and global contexts. Most importantly for the present analyses, they demonstrated that latent semantic analysis (LSA: Landauer & Dumais, 1997) can be used to expose these differences. LSA is both a method for extracting and representing word meanings from a large corpus of text, as well as a theory of knowledge representation (Landauer & Dumais, 1997). LSA provides a useful tool for determining the semantic relatedness between any two units of language (Kurby, Wiemer-Hastings, Ganduri, Magliano, Millis, & McNamara, 2003; Magliano & Millis, 2003; Magliano et al., 2002; Graesser, Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, Person, & the TRG, 1999). In their application of LSA to analyze verbal protocols, they compared the content of a verbal protocol produced at a given sentence to either the current sentence (i.e., the sentence that was read immediately prior to the production of a protocol) or the causally relevant sentences from the prior context, as determined by a causal network analysis (Trabasso, van den Broek, & Suh, 1989). LSA computes a cosine between the mean vector representations of the units of language, which can range from –1 to 1. Magliano et al. (2002), for example, found that sentence-focused sentences had significantly higher cosine with current sentences than both causal sentences and general knowledge of the world (i.e., words that were empirically determined to be related to a sentence); whereas, global strategies had higher combined cosines for causal sentences and general knowledge than the current sentence.

In the current study, LSA was used to assess changes in the strategies as a function of skill and training. We calculated cosines between the pre-SERT and post-SERT self-explanations and three information sources: current sentence, prior text context (i.e., prior two sentences), and the title. The prior text context reflected the local context, whereas the title was a proxy for the global context. That is, readers who were making local connections should explain how a sentence is related to the immediately prior sentence, whereas readers who were connecting the sentence to the global theme should mention information that is closely related to the contents of the title. This approach has been successfully used to expose reading strategies (McNamara et al., in press; Shapiro & McNamara, 2000).

Table 4 contains the cosines for current sentence, prior text, and title. A reading skill by source by SERT mixed ANOVA was conducted in which SERT and source were within-participants variables and reading skill was a between-participants variable. There was a main effect of SERT such that cosines were higher after training ($M = .43$) than before training ($M = .33$), $F(1, 26) = 16.97$ Mse = .02075, $p < .05$. There was also a main effect of source, $F(2, 26) = 116.86$ Mse = .001066, $p < .05$. Post-hoc analyses (LSD) revealed that cosines for the current sentence ($M = .47$) were higher than the prior text ($M = .35$) or title ($M = .32$), which did not differ.

Taken together, these two analyses suggest that although skilled readers become more active after training than less-skilled readers, the self-explanations for both populations become more relevant with the local and global contexts after training. As mentioned above, there is a growing body of research to suggest that individual differences mediate how students respond to reading training because readers use the training to improve skills that are lacking. The current data are consistent with this literature. There was a trend that less-skilled readers had lower strategy scores after training, which would suggest that they emphasized sentence-focused strategies, such as paraphrasing after training. It is also interesting that the less-skilled readers ($M = .56$), had higher cosines with the current sentence than skilled readers ($M = .48$) after training. However, this difference was only marginally significant via a simple *t*-test, $t(26) = 1.72$, $p < .10$.

Table 4. Mean LSA Cosines as a Function of Skill, Sources, and SERT

| | | SERT | |
|---|---|---|---|
| Skill | Source | Pre | Post |
| Skilled | CS | 0.40 (0.15) | 0.48 (0.15) |
| | PT | 0.30 (0.13) | 0.37 (0.11) |
| | Title | 0.29 (0.15) | 0.37 (0.12) |
| Less-skilled | CS | 0.42 (0.13) | 0.56 (0.12) |
| | PT | 0.30 (0.07) | 0.41 (0.07) |
| | Title | 0.25 (0.12) | 0.39 (0.08) |

**Note:** Standard deviations are reported in the parentheses.

*Analysis of True/False Questions*

Table 5 contains the performance on the true/false questions as a function of reading skill, question type, and SERT. A mixed ANOVA was conducted on the proportion of correctly answered true/false questions with Question type (textbase vs simulation model) and SERT (pre vs post). There was a main effect of reading skill such that skilled readers ($M = .85$) answered more questions correctly than less-skilled readers ($M = .78$), $F(1, 26) = 6.36$, Mse = .02204, $p < .05$. There was a main effect of question type such that text-based questions ($M = .87$) were answered correctly more often than bridging questions ($M = .77$), $F(1, 27) = 19.60$, Mse = .01365, $p < .05$. There was also a significant question type by SERT interaction, $F(1, 26) = 5.59$, Mse = .01017, $p < .05$. Post-hoc tests (LSD) revealed that performance on text-based questions did not change as a function of SERT (pre = 88 and post = .86), whereas performance on situation model questions improved as a function of SERT (pre = .74 and post = .80).

These results are consistent with prior research on SERT and other similar strategy training in that this training tends to lead to improvements in the comprehension at the situation model level (O'Reilly et al., 2004). However, this effect is typically seen for high knowledge readers and readers who have strong metacognitive skills. This effect in the current study is encouraging in that the training is designed to improve this aspect of comprehension. For example, a majority of the strategies emphasized in SERT focus on constructing a more coherent and elaborated situation model.

These results are not consistent with prior research suggesting that benefits of training are mediated by individual differences (McNamara, 2004b; O'Reilly et al., 2004). One reason that we may not have found similar results in this experiment was the nature of the comprehension test. Namely the true/false questions may not have been difficult enough to reveal individual differences. Indeed, McNamara has typically used short answer questions to tap

Table 5. Mean Proportion Correct on the True/False Questions as a Function of Question Type, Skill, and SERT

| | | SERT | |
|---|---|---|---|
| Reading skill | Question type | Pre | Post |
| Skilled | Text-based | 0.90 (0.11) | 0.90 (0.10) |
| | Bridging | 0.78 (0.12) | 0.86 (0.12) |
| Less-skilled | Text-based | 0.87 (0.12) | 0.82 (0.12) |
| | Bridging | 0.69 (0.15) | 0.76 (0.11) |

**Note:** Standard deviations are reported in the parentheses.

comprehension as a function of SERT, which are considerably more difficult than the true/false questions used here. As such, Experiment 2 used short-answer questions to assess the effectiveness of strategy training in the context of iSTART.

## EXPERIMENT 2

### Methods

#### *Participants*

Fifty-three students who were enrolled in an introductory psychology course taught at Northern Illinois University participated in pre-training and post-training sessions. Participants received extra credit for participating in the experimental sessions. Three students were excluded from analyses because of high rates of blank responses in both the pre-iSTART and post-iSTART sessions.

#### *Measures of Reading Skill*

As was the case in Experiment 1, the Nelson-Denny reading skills test was used in Experiment 2. However, participants were only given 15 minutes (rather than 20 minutes) to complete the test. This was done to increase the range of performance in the participants (see e.g., McNamara & McDaniel, 2004). The mean score on the Nelson-Denny test for the skilled readers was 22.52. A median split was performed (median = 21.5) to identify skilled and less-skilled readers. The mean score for skilled readers was 27.20 and mean score for the less-skilled readers was 18.19. Performance on the Nelson-Denny test was statistically different between the two groups, $t(48) = 10.31$, $p < .05$.

#### *Task and Design*

The same task and design was employed in Experiment 2 as was used in Experiment 1.

#### *Materials*

The texts on the development of coal and the food chain that were used in Experiment 1 were used in Experiment 2. However, the text on thunderstorm development was used in the practice model of iSTART. As such, a new text was adapted on the movement of glaciers, which was comparable in length (22 sentences) and complexity to the other texts. Furthermore, the text on heart disease was shortened to 20 sentences to match the length of the other texts.

Short answer questions were constructed for Experiment 2. Eight questions were constructed for each text. The answers for text-based questions could be found in 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. As such, 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/the total number of parts of that answer).

### *Coding for Self-Explanations*

The same judge who coded protocols from Experiment 1 coded protocols for Experiment 2, using the criteria and coding procedures.

### *iSTART*

The version of iSTART developed by McNamara et al. (2004) was used in the present study. The iSTART sessions took about an hour and a half.

### *Procedures*

This experiment took place across four sessions within approximately one month. The first session involved introducing the experiment and administering three individual differences tests. One was the Nelson-Denny test, and two assessed domain specific and general science knowledge. Session 1 was run in groups of 20–30 participants. The pre-iSTART and post-iSTART sessions followed the same procedures as Experiment 1. The second session was the pre-iSTART session, which occurred approximately a week after session 1. Participants engaged in iSTART during the third session. They did so at their own pace. They were instructed that they would engage in a computerized reading training system that would help them better comprehend difficult science text and potentially help them become better students. They were also told that the system employed sophisticated technology that could figure out if they were practicing what they learned during the practice session and that they should take practice seriously. The final session was the post-iSTART, which took place one to two weeks after the third session.

## Results and Discussion

Two sets of analyses were conducted. One set involved the self-explanations and the second involved performance on a short-answer comprehension test.

### *Analysis of the Self-Explanations*

The same analyses conducted on the self-explanations from Experiment 1 were conducted on those from Experiment 2. Table 3 contains the mean strategy scores as a function of reading skill and iSTART. A reading skill × iSTART mixed ANOVA was conducted on the strategy scores. There was a significant main effect of iSTART such that the strategy scores were higher post-iSTART ($M = 1.92$) than pre-iSTART ($M = 1.80$), $F(1, 48) = 3{,}86$, Mse = .08150, $p = .055$. A main effect of skill approached significance such that high skilled readers ($M = 1.92$) had a higher strategy score than low skilled readers ($M = 1.78$), $F(1, 48) = 2.98$, Mse = .205, $p < .1$. The interaction between reading skill and iSTART was not significant.

Table 6 contains the cosines for current sentence, prior text, and title. A reading skill by iSTART by source mixed ANOVA was conducted. There was a main effect of iSTART such that cosines were higher after training ($M = .42$) than before training ($M = .35$), $F(1, 48) = 21.64$, Mse = .01339, $p < .05$. There was also a main effect of source, $F(2, 48) = 161.22$, Mse = .001809, $p < .05$. Post-hoc analyses (LSD) revealed that cosines for the current sentence ($M = .50$) were higher than the prior text ($M = .33$) or title ($M = .33$), which did not differ.

These data are consistent with data on live training in that participants had higher strategy scores after iSTART than before iSTART. Furthermore, their self-explanations increased in relevancy with the local and global contexts as a function of training. We did not find evidence of differences between less-skilled and skilled readers as a function of training with respect to their self-explanations.

### *Analysis of Short-Answer Questions*

The average proportions of the questions that were correctly answered were calculated for text-based and bridging questions for each participant. Table 7 contains the means as a function of skill, question type, and iSTART. A Skill by Question type by iSTART mixed repeated measures ANOVA was conducted

Table 6. Mean LSA Cosines as a Function of Skill, Sources, and iSTART

| | | SERT | |
|---|---|---|---|
| Skill | Source | Pre | Post |
| Skilled | CS | 0.47 (0.16) | 0.54 (0.13) |
| | PT | 0.31 (0.09) | 0.36 (0.06) |
| | Title | 0.29 (0.10) | 0.36 (0.06) |
| Less-skilled | CS | 0.46 (0.18) | 0.54 (0.15) |
| | PT | 0.30 (0.10) | 0.35 (0.08) |
| | Title | 0.30 (0.09) | 0.35 (0.09) |

**Note:** Standard deviations are reported in the parentheses.

Table 7. Mean Proportion Correct on the Short-Answer Questions as a Function of Skill, Question Type, and iSTART

| | | iSTART | |
|---|---|---|---|
| Reading skill | Question type | Pre | Post |
| Skilled | Text-based | 0.67 (0.26) | 0.65 (0.22) |
| | Bridging | 0.38 (0.14) | 0.47 (0.15) |
| Less-skilled | Text-based | 0.52 (0.20) | 0.65 (0.19) |
| | Bridging | 0.33 (0.11) | 0.36 (0.12) |

**Note:** Standard deviations are reported in the parentheses.

on the proportion of correctly answered questions. There was a main effect of reading skill such that skilled readers ($M = .56$) answered more questions correctly than less-skilled readers ($M = .47$), $F(1, 48) = 7.15$, Mse = .04954, $p < .05$. There was a main effect of question type such that text-based questions ($M = .63$) were answered correctly more often than situation model questions ($M = .39$), $F(1, 43) = 305.52$, Mse = .009654, $p < .05$. Finally, there was significant Reading Skill by Question type by iSTART interaction, $F(1, 48) = 4.55$, Mse = .02511, $p < .05$. Post-hoc tests (LSD) revealed that less-skilled readers improved in their performance on text-based questions after iSTART, whereas skilled readers did not. On the other hand, skilled readers improved in their ability to answer situation models questions after iSTART, but less-skilled readers did not.

The results of the short-answer questions are consistent with prior research that suggests that different groups of readers benefit in different ways to self-explanation training (McNamara, 2004b; McNamara & Scott, 1999; O'Reilly et al., 2004). For example, McNamara and her colleagues have shown that low knowledge readers develop stronger textbase representations post-SERT and iSTART training (McNamara, 2004b; O'Reilly, Best, & McNamara, 2004; O'Reilly, Sinclair, & McNamara, 2004). The current results suggest that the same is true for computer-based training. We also found that skilled readers increase performance on questions that tap the situation model.

## GENERAL DISCUSSION

The goal of this study was to compare the effectiveness of iSTART to live SERT for skilled and less skilled readers. We were particularly interested in changes in the general reading strategies and relevance of self-explanations as a function of training. With respect to live training, skilled readers engaged in more global processing after training than before, whereas less skilled readers did not. However, both skilled and less-skilled readers produced self-explanations that were more relevant to the local and global contexts after training. What is encouraging about the current study is that we found that both skilled and less-skilled readers increased their strategy scores and produced more relevant self-explanations after iSTART than before. Taken together, these results suggest that both skilled and less-skilled readers change the way they read as a function of both live and computer-based strategy training (see also, O'Reilly, Sinclair, & McNamara, 2004).

With respect to the improvements in comprehension, the current study is consistent with a growing body of research, which suggests that strategy training is effective (e.g., McNamara, 2004b). Students who received SERT answered more bridging questions correctly after training, but not text-based questions. Furthermore, there were no differences as a function of reading skill. These are encouraging results. However, it is possible that the true/false questions used on Experiment 1 were not sensitive enough to tap into these individual differences.

There were, however, individual differences in performance on the short-answer questions as a function of iSTART. Consistent with prior research on live SERT, less skilled readers improved their performance on text-based questions, but not bridging questions. The opposite was found for skilled readers. More importantly for the goals of this study, we demonstrate that computer-based training has similar effects as live reading training.

The results of this study address a critical goal for developing computer-based reading interventions. Specifically, developers of these interventions must demonstrate that they work at least as well as live training. We believe that these results illustrate that computer-based interventions for reading are viable. Computer-based reading training circumvents a major obstacle in providing broadly available reading training. Specifically, teachers may not be willing to devote class time to reading remediation. This is particularly the case for middle school, high school, and early college students, for which this version of iSTART is designed.

Computer-based reading training has at least one possible advantage to live, classroom-based training. Specifically, teachers cannot tailor training to the individual needs of the student when administering live SERT. In contrast, it is quite possible to modify iSTART to the needs of the student. For example, it would be possible to assess the student's reading skill prior to participating in iSTART. Less-skilled readers apparently need remediation in constructing a textbase representation, whereas skilled readers do not. As such, the modules could be adapted to this need of a less-skilled reader. It would be equally possible to have readers interact with iSTART over several iterations and scaffold training as a reader begins to develop new skills and strategies. Indeed, the ongoing modifications of iSTART are working toward more adaptively responding to the students' needs, based both on prior assessments and their on-line performance.

## REFERENCES

Bartlett, F. C. (1932). *Remembering*. Cambridge, England: Cambridge University Press.

Best, R., Ozuru, Y., & McNamara, D. S. (2004). Self-explaining science texts: Strategies, knowledge, and reading skill. In Y. B. Yasmin, W. A. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera (Eds.), *Proceedings of the sixth international conference of the learning sciences: Embracing diversity in the learning sciences* (pp. 89-96). Mahwah, NJ: Erlbaum.

Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *The Nelson-Denny reading test*. Itasca, IL: Riverside Publishing.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, R., & Glaser, R. (1989). Self-explanation: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145-182.

Chi, M. T. H., De Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439-477.

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.

Craig, S. D., Driscoll, D., & Gholson, B. (2004). Constructing knowledge from dialog in an intelligent tutoring system: Interactive learning, vicarious learning, and pedagogical agents. *Journal of Educational Multimedia and Hypermedia*, *13*, 163-183.

Craig, S. D., Gholson, B., & Driscoll, D. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, *94*, 428-434.

Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. *Review of Educational Psychology*, *60*, 517-529.

Graesser, A. C., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Person, N., & the TRG (1999). AutoTutor: A simulation of a human tutor. *Journal of Cognitive Systems Research*, *1*, 31-51.

Graesser, A. C., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Person, N., & the TRG (2000). Using latent semantic analysis to evaluate the contributions of students in AutoTutor. *Interactive Learning Environments*, *8*, 129-148.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371-395.

Knuth, D. (1998). *The art of computer programming,* Vol. 3, 2nd Ed. Reading, MA: Addison-Wesley.

Kurby, C. A., Wiemer-Hastings, K., Ganduri, N., Magliano, J. P., Millis, K. K., & McNamara, D. S. (2003). Computerizing reading training: Evaluation of a latent semantic analysis space for science text. *Behavior Research Methods, Instruments, & Computers*, *35*, 244-250.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211-240.

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure. *Cognition and Instruction*, *21*, 251-283.

Magliano, J. P., Wiemer-Hastings, K., Millis, K. K., Munoz, B. D., & McNamara, D. (2002). Using latent semantic analysis to assess reader strategies. *Behavior Research Methods, Instruments, and Computers*, *34*, 181-188.

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processes during comprehension. *Journal of Educational Psychology*, *91*, 615-629.

McNamara, D. S. (2004a). Aprendiendo a través de los textos: Efectos de la estructura del texto y las estrategias de los lectores. *Signos*, *37*, 123-159.

McNamara, D. S. (2004b). SERT: Self-explanation reading training. *Discourse Processes*, *38*, 1-30.

McNamara, D. S., Best, R., & Castellano, C. (2004). *Learning from text: Facilitating and enhancing comprehension*. www.speechpathology.com

McNamara, D. S., Boonthum, C., Levinstein, I. B., & Millis, K. (In press). Using LSA and word-based measures to assess self-explanations in iSTART. In T. Landauer, D. S. McNamara, S. Dennis & W. Kintsch (Eds.), *LSA: A road to meaning.* Mahwah, NJ: Erlbaum.

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.

McNamara, D. S., & McDaniel, M. (2004). Suppressing irrelevant information: Knowledge activation or inhibition? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 465-482.

McNamara, D. S., & Scott, J. L. (1999). Training reading strategies. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the twenty-first annual meeting of the cognitive science society* (pp. 387-392). Hillsdale, NJ: Erlbaum.

O'Reilly, T., Best, R., & McNamara, D. S. (2004). Self-explanation reading training: Effects for low-knowledge readers. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the twenty-sixth annual meeting of the cognitive science society* (pp, 1053-1058). Mahwah, NJ: Erlbaum.

O'Reilly, T., Sinclair, G., & McNamara, D. S. (2004). Reading strategy training: Automated versus live. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the twenty-sixth annual meeting of the cognitive science society* (pp. 1059-1064). Mahwah, NJ: Erlbaum.

Otero, J. (2002). Noticing and fixing difficulties while understanding science texts. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension*. Mahwah, NJ: Erlbaum.

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 twenty-sixth annual meeting of the cognitive science society* (pp. 1071-1076). Mahwah, NJ: Erlbaum.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. *Cognition and Instruction*, *2*, 117-175.

Shapiro, A., & McNamara, D. (2000). The use of latent semantic analysis as a tool for the quantitative assessment of understanding and knowledge. *Journal of Educational Computing Research*, *22*, 1-36.

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during text comprehension. *Discourse Processes*, *21*, 255-288.

Trabasso, T., van den Broek, P., & Suh, S. (1989). Logical necessity and transitivity of causal relations in the representation of stories. *Discourse Processes*, *12*, 1-25.

van den Broek, R., Risden, K., & Husebye-Hartman, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In R. F. Lorch & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 353-374). Hillsdale, NJ: Erlbaum.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies in discourse comprehension.* New York: Academic Press.

Yuill, N. M., & Oakhill, J. V. (1988). Effects of inference awareness training on poor reading comprehension. *Applied Cognitive Psychology, 2,* 33-45.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123,* 162-185.

Direct reprint requests to:

Joe Magliano
Department of Psychology
Northern Illinois University
Dekalb, IL 60115
e-mail: jmagliano@niu.edu