Progress Summary

Promoting Active Reading Strategies to Improve Students' Understanding of Science

Danielle S. McNamara

Start Date: Oct 1 2000
Expires: Sep 30 2005

The problem addressed in this five-year project regards high-school students' understanding of and ability to learn from difficult science texts. The study is designed to address several issues. The first goal of this project is to examine high school students' and teachers' knowledge about metacognitive reading strategies to better ascertain the need for strategy interventions, as well as teachers' ability to provide such training to students. The second goal is to compare the effectiveness of three reading interventions and to thereby identify more effective methods for improving students' comprehension skills. The third goal is to determine whether the benefits of reading interventions depend on students' skills, knowledge, or cultural environment. The fourth goal is to develop a computerized training program that adaptively assigns a student to an appropriate training regime that interactively provides reading training.

Goal 1: Examine students' and teachers' knowledge about metacognitive reading strategies to better ascertain the need for strategy interventions.

Metacognition refers to an individual's monitoring of cognitive processes and the knowledge and use of cognitive processes involved in successful learning (Hacker, Dunlosky, & Graesser, 1998; Reutzel & Cooter, 2000; Tei & Stewart, 1985). Readers better understand and learn more from written material when they monitor their comprehension and use active reading strategies such as previewing, predicting, making inferences, drawing from background knowledge, and summarizing. Unfortunately, there is relatively little research concerning high school students' knowledge of and use of strategies, its impact on academic performance, and how that relates to their teachers' knowledge of reading strategies and classroom instruction.

To address our first goal, we have completed three experiments. In the first study (O'Reilly, McNamara, and the Strategies Lab, 2002), 1651 students from four school locales (rural, rural Appalachian, suburban, and inner city) were measured in terms of their reading skill, prior science knowledge, and knowledge of metacognitive reading strategies. Science achievement was assessed in terms of the students' understanding of a science passage, course grade, and a state standards-of-learning (SOL) test. Concurrent with our hypotheses, we found that students' knowledge, reading skill and metacognitive reading strategies play an important role in determining the students' comprehension of scientific texts and their science course grades. More importantly, the findings underscored the significance of metacognition in learning. Students with greater knowledge of metacognitive reading strategies scored higher on measures of science passage comprehension and state standards-of-learning tests, and had higher course grades than students with less knowledge of metacognitive reading strategies. Knowledge of metacognitive reading strategies was also positively related to the number of books read, positive attitudes towards reading and learning science, and the amount of time and effort the students put into their courses. Even more encouraging is the finding that students with a high level of reading strategy knowledge were able to compensate for a meager knowledge base. While prior knowledge is critical for comprehension, it appears that metacognitive reading strategies can help students overcome difficulties in acquiring new material even when they have limited knowledge of novel information.

The second study to assess the relationships between knowledge, reading skill, knowledge of reading strategies, and course performance was conducted with 144 college students enrolled in an introductory psychology course (Cottrell & McNamara, 2002). This study was conducted to examine whether the same relationships between course performance and individual differences were found with an older and more skilled population. In addition, it was considered that the college students' performance would be reflective of what one might find for high-school teachers. We examined the students' prior knowledge, reading ability, reading strategy knowledge, science text comprehension, and course grades. Prior knowledge and reading skill were predictive of science text comprehension, but only prior knowledge predicted course performance. However, prior knowledge did not benefit students who did not use reading strategies. In contrast to first study with high-school students, reading strategy knowledge did not compensate for less domain knowledge. This result is potentially due to the age and skill advantages in comparison to high-school students. However, this result may also be due to the measure used to assess reading strategy knowledge. Future studies will investigate these issues.

The third study more directly addresses the question regarding the relationship between teacher practices and student performance. Thus far, 40 teachers have been administered two surveys to assess their attitudes and beliefs, instructional practices, and knowledge of reading strategies. Their students (n=3000) have been assessed in terms of their science knowledge, reading skills, reading strategies, and course performance. These data are currently being compiled and analyzed.

Goals 2 and 3: Compare the effectiveness of reading interventions and assess whether the benefits of these reading interventions depend on students' skills, knowledge, or cultural environment.

We are comparing the following three programs in terms of increasing high-school students' comprehension of science texts: Self-Explanation Reading Training (SERT), Previewing (PRE), and the Interactive Notation System for Effective Reading and Thinking (INSERT). SERT is based on research demonstrating that when students self explain difficult text aloud, their understanding of the reading material is enhanced (Chi & Bassok, 1989; McNamara & Scott, 1999). Previewing, or quickly reading over important parts of a text before reading it, has been shown to improve high-school students' reading comprehension (Golembesky, Bean, & Goldstein, 1997; Johnston & Allington, 1991; Richardson & Morgan, 2000). Previewing involves skimming a text before reading it and focusing on parts of the text such as key words, headers, the table of contents, summaries, and so on. The INSERT strategy instructs readers to use a notation system to record symbols next to each paragraph to indicate agreement, importance, novelty, and understanding (see Estes & Vaughan, 1986; Singer & Donlon, 1985).

These notations are intended to cue the reader to think on higher levels of cognition about what is being gained while reading the text.

Our first experiment included 339 students from 16 ninth-grade Earth Science classrooms across two schools in Chesapeake, Virginia. The results indicated that students who were provided with strategy instruction scored higher on measures of science comprehension and learning than students who were not given such instruction. However, not all forms of strategy interventions equally benefited students' comprehension. Students who were provided with SERT training consistently performed better than students who were trained on strategies that emphasized either prediction and prior knowledge activation (Preview strategy) or on-line comprehension monitoring (INSERT). These results indicate that active, explanation based processing of written material is crucial to understanding. Self- explanation training insures that the student learns how to make gap-filling inferences when reading difficult scientific texts.

These issues are currently being investigated on a larger scale, including approximately 1500 high-school students from three locales (rural Appalachian, suburban, and inner city). During the past school year, approximately 70 Physical Science, Earth Science, Biology, and Chemistry classrooms have participated. Students were administered pretests at the beginning of the school year to assess science knowledge, reading skill, reading strategy knowledge, and science text comprehension. In the middle of the school year, students were administered one of the reading strategy interventions. They were again tested in terms of science comprehension approximately one month following training and at the end of the school year. These data are currently being compiled and analyses will be conducted during the upcoming academic semester.

It was observed in our first experiment that the comprehension strategy interventions were not effective for students with very low reading skills. Therefore, a fourth strategy intervention has been developed to provide instruction regarding word comprehension. This intervention focuses on strategies such as breaking the word into parts, using the sentence and pictures to figure out the meaning, INSERT. We compared the effectiveness of this intervention to SERT-plus (a combination of SERT and Previewing) and a Control group across nine Earth Science classrooms. This data will also be analyzed in the 2002-2003 academic year.

As might be expected, numerous challenges were encountered in our large-scale intervention study. The most significant problems in terms of scaling up were low student class attendance and attrition (most significantly for the most needy students) and unreliable teacher cooperation. In addition, student dynamics often interfered with strategy practice during training. Our expectation is that the automated tutor has the potential to resolve many of the problems encountered during this 'live training' phase, though undoubtedly we will encounter others of a different nature.

Goal 4: To develop a computerized tutor that provides reading training, practice, and feedback.

Finally, we have made substantial progress toward the goal of developing our automated program, Interactive Strategy Training for Active Reading and Thinking (ISTART). On-screen characters (created using Microsoft Agent®) provide instruction regarding self-explanation and supportive reading strategies such as comprehension monitoring, making bridging inferences, and using logic to make sense of the text. Following a demonstration of the strategies by the agents, the students then practice the strategies by reading a text and typing their self-explanations. During this phase, the agents' interactions with the student are moderated by the quality of the explanation (e.g., increased interactions and support with shorter, less relevant explanations). We recently debuted the ISTART program during the Spring, 2002 semester in an experiment with 300 college students enrolled in an introductory biology course. One third of the students were given live SERT training, and one-third were given training using the ISTART program. We have not yet compiled the experimental results in terms of impact on science comprehension and course performance. However, our preliminary results indicate that the ISTART program's judgment of students' self-explanation quality was surprisingly accurate. The success of the ISTART program was further evidenced by positive student reactions.

The ISTART program was also tested with 20 eighth and 20 ninth grade students in June, 2002. Half of the students from each grade level were assigned to a control condition and half to the training condition. The control condition were provided with a brief instruction concerning self-explanation and then read a text about heart disease and typed their self-explanation after reading each sentence. The training group was provided with self-explanation training before reading and self-explaining the heart disease text. The students' reactions to the ISTART program were positive, indicating that they enjoyed working with the program and that they valued what they learned.

Our second aim to achieve the goal of developing ISTART was to integrate latent semantic analysis (LSA) into the computer program to improve its ability to interpret the students' input. There have been three objectives to achieve this objective. The first is collecting and analyzing protocols before and after students receive SERT training. The second is further examining ways in which LSA values can identify SERT strategies in a tutoring environment. The third goal is building a LSA database that maximizes success in this regard.

In regard to the first objective, two experiments have been conducted. In the first study, college students were asked to self-explain texts before and 1-2 weeks after receiving SERT (live) training. We classified the protocols on the extent they elaborated the text: minimal (none), text-focused (a little), and knowledge-building (added elaborations from world knowledge). Across sessions, better readers decreased their proportion of minimalist responses and increased their text-focused and knowledge-building responses. Poorer readers showed the opposite pattern: they decreased sentence-focused and knowledge-building responses and increased their minimalist responses. We are currently examining why the poorer readers showed this counter-intuitive pattern. One potential reason is that minimalist responses are typically paraphrases, and paraphrasing is one of the first strategies taught and it is one of the easiest to do. Perhaps poorer readers need more practice on SERT before they can move on to knowledge-building or they need feedback during practice sessions for them to do more elaborative self-explanations. Alternatively, the minimalist responses may have included logic and common sense that was missed within the large grained scoring system. We are currently examining these possibilities.

A second study was conducted with two delays between SERT training and post-SERT assessment: short delay (1 – 2 weeks) or long delay (6-7 weeks). We also included a test of prior knowledge to examine whether reading ability and prior knowledge interact in terms of SERT effectiveness. Furthermore, we added a short-answer comprehension test in addition to a T/F test used in the first study. We are currently analyzing these data.

In regard to the second objective, we have been continuing work on using LSA to assess strategies in a number of ways (Magliano, Wiemer-Hastings, Millis, Muñoz, & McNamara, D.S. in press; Millis, Magliano, Wiemer-Hastings, & McNamara, in press). We have found that LSA successfully identifies paraphrases (minimalist responses) and detailed self-explanations (knowledge-building). We intend to further examine whether LSA can identify particular ideas, such as a particular bridging inference or elaboration.

In terms of the third objective of building the LSA database, we have made the following progress. A corpus of texts covering relevant sciences (Physics, Chemistry, Biology, Earth Sciences) was developed to train a topic-specific LSA space for the SERT trainer. The expectation is that an LSA space based on text that is relevant to the texts that students read and write about will allow for more accurate LSA judgments (cosines) of student contributions relative to the comparison texts, which we call "benchmarks".

To find the best performing LSA space, several aspects of the text and parameters of the LSA training were varied. The text characteristics were (i) the *amount* of text, (ii) the *proportion of text* about *specific* topics (e.g., specific articles) as opposed to *general* information (e.g., textbook chapter introductions), and (iii) the *format* of the text as it was read into the LSA, i.e., a sentence-by-sentence format versus a paragraph-by-paragraph format. We further varied the *number of dimensions* when training the LSA. The number of dimensions can have striking effects on how well LSA cosines match human judgments of text similarity.

To evaluate the LSA spaces resulting from the combination of above factors, a set of 90 student think-aloud contributions were selected that varied on the level of processing that the students were engaging in (e.g., minimalist versus knowledge-building). Four human raters compared each statement to five benchmarks to provide a standard for comparing LSA (thus, 450 judgments were made altogether). These benchmarks included a list of words representing the current sentence, the prior text, world knowledge, and two general reading strategies. Two of the human raters were expert text researchers (expert ratings)and two others were trained students (novice ratings). To evaluate LSA, the cosines obtained from the LSA spaces for each pair of student contribution and benchmark was compared to the ratings provided by humans. Reliable LSA judgments of these texts are reflected by significant correlations between the LSA cosines and the human ratings.

The text corpus that produced the best match with the human ratings contained all of the available text and was formatted in paragraphs. The best dimensions were 250 and 450. Previous analyses utilizing LSA performed in the context of this grant made use of an LSA space that is available online (http://lsa.colorado.edu). The corpus used for all of these was the General Reading (LSA-GR) corpus, which contains an extensive amount of text. A critical test of our designated LSA space is to compare the correlations to those obtained with cosines of this General Reading LSA space to those obtained with the science text corpuses that we developed.

Ideally, correlations will be higher with our science textbase because this is a domain-specific semantic space that should be better attuned to the content of the student protocols.

The 90 sentences were submitted to the LSA-GR, compared against the 5 benchmarks. LSA-GR correlated significantly with all human raters ($p<0.001$). The correlations varied from 0.50 to 0.54 (individual raters). The correlation with the average expert rating and the average novice rating was 0.55. In comparison, the cosines obtained from our own corpus trained on 450 dimensions yielded higher correlations with human raters: The correlation with the averaged expert ratings was 0.72 ($p < 0.001$); cosines correlated 0.69 ($p < 0.001$) with averaged novice ratings. The individual human –LSA correlations varied from 0.59 ($p < 0.001$) for one novice rater to 0.68 ($p < 0.001$) for the remaining raters, i.e., one novice and two expert raters. These correlations are excellent both in comparison to correlations obtained in other projects, as well as in comparison to the correlation achieved between the human raters, which varied from 0.69 ($p < 0.001$) to 0.80 ($p < 0.001$). These results show that our newly developed LSA space performs reliably and better than the previously used LSA space.

Work to be completed

In conclusion, we have made substantial progress to achieve our goals. Year 3 will be devoted primarily to the analyses and interpretation of the large amounts of data collected during the first two years of this project. In addition, we will continue our work to develop and refine ISTART. The development of ISTART will be informed by the research we have conducted thus far such that the students' interactions with the computer program can be modified according to student ability. Year 4 will be devoted to testing the effectiveness of ISTART in classroom settings and examining the level of support needed by the teachers to use the system (please see original proposal). Year 5 will be devoted to the analysis of this data and further development of the system based on our outcomes.

REFERENCES

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.

Cottrell, K., & McNamara, D.S. (in press). Cognitive precursors to science comprehension. *Proceedings of the Twenty-fourth Annual Meeting of the Cognitive Science Society*.

Golembesky, B. L. Bean, R. M., & Goldstein, H. (1997). Effects of congruence in Title I reading using PUSH-UPS: A story previewing technique. *Reading Research and Instruction, 37,* 1, pp. 61-83.

Graesser, A., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Person, N., & the Tutoring Research Group, (in press). Using Latent Semantic Analysis to Evaluate the Contributions of Students in AutoTutor. To appear in a special issue of *Interactive Learning Environments*, J. Psotka, guest editor.

Graesser, A. C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R., & the Tutoring Research Group, (in press). (in press). AutoTutor: A simulation of a human tutor. *Journal of Cognitive Systems Research*.

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (1998). *Metacognition in educational theory and practice*. Mawhwah, NJ: Erlbaum

Johnston, P., & Allington, R. (1991). Remediation. In P. D. Pearson, R. Barr, M. L. Kamil, P. Mosenthal (Eds.), *Handbook of reading research* (pp. 984-1012). White Plains, NY: Longman Inc.

Magliano, J.P., Wiemer-Hastings, K. Millis, K. K., Muñoz, B.D., & McNamara, D.S. (in press). Using latent semantic analysis to assess reader strategies. *Behavior Research Methods, Instruments, and Computers*.

Millis, K.K, Magliano, J.P., Wiemer-Hastings, K., & McNamara, D.S. (2001). Using LSA in a computer-based test of reading comprehension. In J.D. Moore, C. Luckhardt-Redfield, & W.L. Johnson (Eds.), *Artificial intelligence in education: AI-ED in the wired and wireless future: Vol. 68. Frontiers in artificial intelligence and applications* (pp. 583-585). Amsterdam, The Netherlands: IOS Press.

McNamara, D. S., & Scott, J. L. (1999). Training reading strategies. *Proceedings of the Twenty- first Annual Meeting of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

O'Reilly, T., & McNamara, D.S. (in press). What's a science student to do? *Proceedings of the Twenty-fourth Annual Meeting of the Cognitive Science Society*.

Reutzel, D. R., & Cooter, R. B. Jr. (2000). *Teaching children to read.* 3rd Edition. Columbus, Oh: Merrill.

Singer, H. & Donlon, D. (1985). *Reading and learning from text.* Hillsdale, N. J. : Erlbaum.

Tei, E. & Stewart, O. (1985). Effective studying from text. *Forum for reading, 16*, 2, 46-55.

Vaughan, J. L. & Estes, T. H. (1986). *Reading and reasoning beyond the primary grades*. Boston, MA: Allyn and Bacon, Inc.