Classroom Based Reading Strategy Training:
Self-Explanation vs. a Reading Control

Tenaha O'Reilly (t.oreilly@mail.psyc.memphis.edu)
Roger S. Taylor (rtaylor@mail.psyc.memphis.edu)
Danielle S. McNamara (d.mcnamara@mail.psyc.memphis.edu)

Institute for Intelligent Systems
Psychology Department, University of Memphis
Memphis, TN 38152 USA

Abstract

This study examined the effects of a reading intervention called Self-Explanation Reading Strategy Training, (SERT) on high-school students' comprehension of science text. Students ($n = 465$) in 19 classrooms from three high schools were randomly assigned to either SERT or a Control condition. Science comprehension was assessed immediately after training with a science passage about the origin of viruses. The results indicated that participants who mastered SERT strategies outperformed control participants. However, further analyses revealed that the effects of condition depended upon the school in which the students were tested. Implications for individualized training are discussed.

Introduction

Today's high school classrooms have become increasingly diverse in terms of both culture and student ability level (Stormont, 2005). Diversity poses some difficult challenges for educators, as teachers are faced with accommodating a wide range of student backgrounds and knowledge levels (Stormont, 2005). Moreover, teachers are expected to teach a wide range of students in terms of ability, such as reading skill. As a consequence, many educators teach in a way that accommodates the average student (Fuchs, Fuchs, Mathes, & Simmons., 1997). However, the outcome of using this form of accommodation may do little to help students who are not performing at the average level - the low or high ability students.

While extensive research has shown that smaller class sizes significantly improve learning (Finn, Gerber, Achilles, & Boyd-Zaharias, 2001), it is often not possible to reduce class sizes because of limited funding. Thus, in regular classes it is often difficult for teachers to provide the individualized instruction students may need.

One approach to help teachers cope with the challenges of diversity is to teach general leaning strategies. Teaching learning strategies helps students improve self-efficacy (Van-Keer & Verhaeghe, 2005) and autonomy (Rivers, 2001). More importantly, teaching learning strategies, such as those directed at improving reading comprehension has been shown to be effective (Rosenshine & Meister, 1994; Rosenshine, Meister, & Chapman, 1996).

One such reading intervention that has gained attention in recent years is self-explanation (Chi, De Leeuw, Chiu, & LaVancher, 1994). Self-explanation is the process of explaining the meaning of text to one's self while reading. Research has shown that training students to self-explain improves learning and comprehension (Chi et al., 1994).

More recently, McNamara (2004) has developed a reading strategy intervention called Self-Explanation Reading Strategy Training (SERT) based upon the concept of self-explanation. SERT builds upon previous self-explanation research by integrating a variety of empirically based reading strategies to scaffold students' ability to self-explain. The training progresses from teaching relatively simple techniques designed to help local processing such as paraphrasing to more demanding strategies that require global processing such as bridging and elaboration. Research has indicated that SERT helps improve reading comprehension in laboratory studies with college students (McNamara, 2004) and middle school students (McNamara, O'Reilly, Best, & Ozuru , in press), as well as in a small-scale classroom study with high school students (O'Reilly, Best, & McNamara, 2004).

These results are promising; however, the past studies have been limited in two ways. First, the majority of them have been conducted in the laboratory rather than in classrooms, and research has shown that the effects produced in the lab may not generalize to real world settings (e.g., Ceci & Bronfenbrenner, 1985). For example, Ceci and Bronfenbrenner (1985) found that children's (10 and 14 year olds) use of time monitoring strategies differed as a function of the location of the experiment. Children's use of time monitoring strategies was less strategic in the lab than when the children were tested in a natural setting. Thus, it is important to assess whether strategy training based on lab results would be effective in an authentic school setting. For example, in some of the early lab based research (e.g., McNamara, 2004) participants were trained in strategy use on a "one-on-one" basis. In contrast, in the current study, participants were trained during their regular class time, and thus the training was given to the class as a whole. In real classroom settings, extraneous variables such as disruptions, peer interactions, teacher attitudes, and attendance may mitigate the impact of the intervention.

A second limitation of previous research was that the scale of the studies has been relatively small. Our goal here was to

more broadly examine the impact of strategy training for a larger number of high school students across diverse settings and from a variety of socio-cultural backgrounds. This type of research will help to determine whether the strategy training can be scaled up to a wide variety of real-world settings.

In the current study, we examined the effect of SERT and a reading control on high school students' comprehension of a science text. The study included a large sample of students (*n* = 465) from three different high schools in two states (i.e., Virginia and Kentucky). Students were also tested in a variety of science classes. These factors help to provide a diverse sample to assess the scalability of the strategies.

We predicted that students in the SERT condition would outperform students in the control condition because SERT was designed to help students to utilize their prior knowledge, which research has shown to be critical to reading comprehension (Shapiro, 2004).In addition, SERT should help students to build an integrated text representation that incorporates both local and global text processing.

SERT requires the student to learn new skills, so they must first learn the new declarative information, and then gradually proceduralize this information over time (Anderson & Lebiere, 1998). Thus one possibility is that the students in the SERT condition may experience cognitive overload during the training, which may actually impair the comprehension in comparison to the control condition. Students in a reading control condition may use reading techniques with which they are already familiar and proficient at using. These techniques are likely to be relatively automatized and, as a result, they may interfere less with the reading and learning process. Because SERT is a new and potentially more cognitively demanding task, there is a possibility that some students may perform more poorly on the post-training comprehension measures in comparison to the students in the more naturalistic control condition.

Method

Nineteen classrooms in three schools were randomly assigned to the SERT or control conditions. Students in the SERT classrooms were provided with training on how to self-explain texts using five reading strategies. Control classrooms were simply asked to read the same science text as was used with the SERT training. Comprehension was assessed using a new science passage taken from a high school science textbook. Two question formats were used to evaluate the students' understanding: *multiple choice* and *open-ended questions*. For each question format, we used both text-based and bridging-inference questions to assess different levels of students' understanding of the science passages.

Participants

The sample consisted of 465 students in grades nine to twelve from a variety of science classes including *biology, earth science, physical science, and chemistry*. There was an approximately even distribution of males (50.8%) and females (48.6%) in the sample (0.6% not specified). The

ethnicity was distributed in the following manner: White 57%, African American 24.5%, Native American 5.6%, Hispanic 4.7%, Asian 2.6%, 0.2% not specified. The mean age was 15.3 (*SD* = 1.17). One high school was located in an *urban* region of Norfolk, Virginia, the second in a *suburban* region in Williamsburg, Virginia, and the third in a *rural* region in Prestonsburg, Kentucky.

Materials

Student abilities were measured with two tests: the Gates-MacGinitie Reading Skill Test and a Prior Science Knowledge Test. The Gates-MacGinitie test is a standardized reading comprehension test, designed for grades 10-12. The test consists of 48 multiple-choice questions designed to assess student comprehension on several short text passages (Cronbach's Alpha α=.91). Due to time constraints, the time limit for the comprehension question section was reduced to 15 minutes. The prior knowledge test consisted of 35 multiple-choice items which tap knowledge of different science domains including biology, scientific methods, mathematics, earth science, physics, mathematics, and chemistry (Cronbach's Alpha α=.82).

To examine reading comprehension, immediately following training, we administered a passage on the science topic of *viruses*. The text described the structure, and reproduction of viruses, as well as providing some examples of viruses and how they relate to disease. The passage was 1,216 words in length with a Flesch Reading Ease of 45.1 and Flesch-Kincaid Grade level of 10.6. Reading comprehension was assessed with a set of eight open-ended and eight multiple-choice questions, half were text-based and half were bridging-inference questions. The answer to the text-based questions could be found in a single sentence of the passage. In contrast, the answers to the bridging-inference questions required the reader to integrate the information across two or more sentences in the text. For the purposes of this paper, we analyzed the overall total proportion correct on the virus passage.

Self-Explanation Reading Training (SERT)

The training was delivered in three main phases: introduction, demonstration, and practice (see, McNamara, 2004). During the introduction phase, participants were provided with a description and examples of self-explanation. The instructor defined and provided examples for five reading strategies: comprehension monitoring, paraphrasing, elaboration, logic and common sense, prediction, and making bridging inferences.

During the demonstration phase, participants watched a video depicting a student reading and self-explaining a text about forest fires. Participants could refer to the accompanying video transcript during viewing. The video was paused at predetermined points, and participants identified and discussed the strategies being used by the student in the video. In the practice phase, the participants worked in pairs to practice self-explanation while reading a chapter from their science textbook. The participants took

turns self-explaining, alternating after each paragraph. At the end of each paragraph, the partner who was listening (and not self-explaining) summarized the paragraph.

Participants in the reading control condition were asked to read the same science text on forest fires that the SERT participants had self-explained.

Design and Procedure

Each class was randomly assigned to one of two conditions: SERT or Control. Experimental sessions were conducted by two experimenters during students' regular classroom time. Prior to training, students completed the prior knowledge test followed by the Gates MacGinitie Reading skill test. A 15-minute time limit was given for each test.

SERT training was conducted during two class periods conducted on consecutive days. Participants were told that the purpose of the study was to learn strategies that would help them to better understand and remember what they read. The control condition required one class period. Students in the control condition were told that the purpose of the study was to learn what strategies students use when reading their textbooks. They read the science texts to which trained groups were exposed during training.

Immediately after training, participants were given 30 minutes to read a new science text passage and answer the accompanying comprehension questions. In the SERT condition, the experimenter briefly reviewed the strategies for SERT before beginning the comprehension test. The students did not have the text available when answering the questions.

Results

Pretest Individual Differences Scores

To examine potential pre-training differences in student ability scores, we performed analyses on the students' pretest reading ability and prior science knowledge as a function of condition and school. An analysis of the reading skill scores indicated that there was no significant effect of condition, $F(1, 459) = 1.87$ $MSE = 66.76$, $p = .172$, a significant effect of school $F(2, 459) = 27.42$, $MSE = 66.76$, $p < .001$, and no interaction $F(2, 459) < 1$.. A post hoc least significant difference (LSD) test indicated that the students in the rural school ($M = 19.78$, $SD = 8.23$) scored significantly higher than the students in the urban school ($M = 16.88$, $SD = 8.17$), and students in the suburban school ($M = 23.76$, $SD = 8.17$) scored higher than students in both the urban and rural schools.

The analysis of the prior knowledge scores revealed no effect of condition $F(1, 459) = 1.05$, $MSE = .029$, $p = .736$, a significant effect of school, $F(2, 459) = 26.47$, $MSE = .029$, $p < .001$, and no interaction $F(2, 459) < 1$). A post hoc LSD test revealed that the students in the rural school ($M = .56$, $SD = .17$) scored significantly higher than the students in the urban school ($M = .50$, $SD = .17$), and students in the suburban school ($M = .64$, $SD = .17$) scored higher than students in both the urban and rural schools.

In short, any effects of condition can not be accounted for by pre-training differences in reading ability or science knowledge. However, the results indicated that the schools varied in terms of reading skill and prior knowledge levels. This was expected as we had chosen the schools such that they would provide students with a wide range of skill and knowledge levels.

Overall analysis: Effects of Training Condition

A 2 X 3 ANOVA was conducted on the total proportion correct on the science text passage (virus) as a function of training condition (SERT, Control) and school (urban, suburban, and rural). To control for the effects of individual differences, we entered science knowledge and reading skill as covariates.

Neither the effect of condition, $F(1, 457) = < 1$, school, $F(2, 457) = 2.05$, MSE=.021, $p = .130$, nor the interaction between condition and school $F(2, 457) = 1.91$ $MSE = .041$, $p = .149$, were significant.

Criterion-Based Performance

One potential reason for the lack of a training effect is that some of the students may have acquired an incomplete understanding of the SERT strategy. That is, there may be large differences in terms of acquiring strategy proficiency. Thus, the lack of a clear difference between the control and SERT conditions may be attributable to inadequate mastery of the SERT reading strategies by some students. This result would indicate that SERT was potentially effective, but that additional training may be required for some students to master the strategies.

To explore this possibility, we categorized students based on their level of mastery of the concepts explained in SERT. During training, SERT participants were given a 12-item multiple-choice SERT quiz that assessed their knowledge of the SERT strategies. We set a criterion of mastery relatively leniently at 66% correct on the SERT test. According to this criterion, 93 (42%) of the total number of participants trained to use SERT met this criterion of mastery, whereas 103 students (46%) did not. Twenty-seven of the SERT participants (12% of the SERT sample) were not included in the analysis because they did not complete the SERT mastery test.

Two 2 X 3 between-subjects ANOVA's with condition (SERT vs. Control) and school as factors were performed on the comprehension questions (knowledge and reading skill were entered as covariates). The first ANOVA included the SERT participants who did *not* meet our criterion of mastery. For non-mastery participants, there was no effect of condition $F(1, 337) < 1$, and no significant interaction of condition and school, $F(1, 337) = 1.34$, $MSE = .027$, $p = .262$.

In contrast, the analysis with mastery students revealed a significant effect of condition $F(1, 337) = 5.41$, $MSE = .112$, $p = .02$, Cohen's $d = .24$, indicating that SERT participants scored higher ($M = .62$, $SD = .17$) than control participants ($M = .58$, $SD = .16$). However, there was also a significant

interaction between condition and school, $F(2, 327) = 4.59$, $MSE = .021$, $p = .01$.

Follow-up analyses were conducted to analyze this effect. For students in the rural school, there was a significant advantage for the SERT group ($M = .62$, $SD = .14$) over the control group ($M = .52$, $SD = .14$), $F(1, 99) = 6.50$, $MSE = .019$, $p = .01$, Cohen's $d = .71$. For the urban school, there was a trend for SERT ($M = .57$, $SD = .15$) to score higher than control ($M = .52$, $SD = .14$), but the result was not significant, $F(1, 69) = 1.75$, $MSE = .021$, $p = .19$, Cohen's $d = .34$. For the suburban school there was no significant difference between the SERT ($M = .62$, $SD = .15$) and the control ($M = .64$, $SD = .15$) conditions, $F(1, 155) < 1$, Cohen's $d = .13$.

Mastery as a Function of School

In sum, our results revealed that SERT can improve students' comprehension of scientific text. However, the findings also indicated that the effectiveness of the intervention depended upon the school in which the students were enrolled. This result begs the question as to why SERT was effective in one school but not the others.

Of course, as intended, the schools differed quantitatively in terms of the students' performance on the prior ability measures. As such, it is notable that the school with the highest performance on prior abilities (i.e., suburban) showed no gains, and the school with the lowest performance (i.e., urban) showed only moderate gains. This result thus raises the possibility that the effectiveness of the training depends on students' prior abilities. Quite simply, the students who needed the training but had enough skills to tackle higher level reading comprehension strategies (i.e., rural) were the ones who gained the most from the training.

If prior abilities are responsible for determining who can benefit from the training, then there should be a difference on the SERT scores as a function of school. If students in the suburban school (i.e., higher performing school) did not benefit from the training because they have already mastered the strategies, then they should have the highest scores on the SERT mastery quiz. To test this hypothesis, we performed an analysis of variance on the students' SERT quiz scores as a function of school. Differences among the schools in terms of the SERT quiz scores (i.e., degree of strategy mastery) may reflect why students in some schools benefited from the training while students in the other schools did not.

We conducted an analysis on the SERT quiz scores for the 93 students who met our definition of mastery (scored at least 66% correct on the SERT quiz) as a function of school. An analysis that included all SERT subjects (both mastery *and* non-mastery, $n = 196$) produced identical results.

A three-way ANOVA with school as the between-subjects factor (knowledge and reading skill were entered as covariates) and the SERT quiz score as the dependent variable revealed a significant effect of school, $F(2, 88) = 11.23$, $MSE = .004$, $p < .001$. Follow up analyses revealed that suburban school, participants scored significantly higher ($M = .82$, $SD = .07$) on the SERT quiz than students in the

rural ($M = .75$, $SD = .06$), $F(1, 74) = 16.16$, $MSE = .004$, $p < .001$, Cohen's $d = 1.07$., and the urban school ($M = .76$, $SD = .06$), $F(1, 71) = 8.67$, $MSE = .04$, $p = .004$, Cohen's $d = .92$. There was no significant difference in the SERT quiz scores between the rural and urban schools, $F(1, 29) < 1$.

In sum, our analysis revealed that the effectiveness of the training depended upon school. However, further analyses revealed this difference was mediated by the students' ability to master the strategies: Students in the suburban schools were more likely to master the strategies than students in other schools. One potential explanation for why the students in the suburban school did not show any effects of training is because they may have already known the strategies before training. Moreover, students in the urban school may not have had a large effect of training because they lacked the minimum skills required to learn the strategies.

Discussion

The principal aim of this study was to examine whether SERT could help students better comprehend science text. The design of the study has some notable strengths. First, the training took place in actual classrooms. Second, the sample size was both large and diverse in terms of school locale, student abilities, course type, gender, and ethnicity - the combination of which strengthens the ecological validity of the research.

We predicted that the SERT condition would outperform the control condition because SERT provides students with a variety of strategies that can be used to foster the construction of a more integrated representation of the text. However, comparing the performance of SERT to a reading control condition is a challenging contrast.

It can be argued that our reading control condition was not a strategy-free situation. Students may have used a variety of strategies that they have used successfully for years. In this way, our control participants may have used a collection of strategies, as opposed to using no strategies at all. This possibility is supported by another reading comprehension study (O'Reilly, Sinclair, & McNamara, 2004), in which the researchers asked participants in a control reading condition what strategies they had used while reading the text. Only 16.5% of the participants indicated that they had not used any strategies. In fact, control participants reported using a variety of strategies that included identifying key points 14.7%, mnemonics 6.4%, imagery 7.4%, and using prior knowledge 7.3%. In short, it is possible that the participants in our control condition utilized other effective strategies.

Despite this possibility, we were still surprised to find that, at first glance, the SERT condition was not statistically different from the control condition. One interpretation of this finding is that students might not have adequately learned the SERT strategies. To test this theory, we analyzed the participants in terms of their scores on a test of SERT strategy mastery. The analysis revealed that for those students who *did not* master SERT, there was no difference in their performance as compared to those in the control condition. In

contrast, for the participants who *did master* the strategies, there was a large advantage of SERT over the control condition (Cohen's *d* = .71 in the rural school).

While these results were encouraging, the impact of the training depended upon the school in which the strategy was tested. There was a large and clear effect in the rural school. However, in the urban school we only found a trend of an effect, and in the suburban school, there was no effect. Qualitatively and quantitatively, we can speculate as to why we did not get a uniform effect in all schools.

Qualitatively, we can say that the schools differed in many ways such as location, socio-economic status, and ethnic composition. However, the precise manner in which these factors affected the outcomes of this experiment is unclear. Nonetheless, our results show that the impact of strategy intervention in the context of real classroom situations is a complex phenomenon. Researchers need to be aware that because a certain intervention does, or does not work, in one context, doesn't necessarily mean that the intervention is effective in other contexts.

Quantitatively, our results have shown that all the schools differed on both our measures of ability. Students in the suburban school had highest reading skill and knowledge scores followed by the students in the rural and urban schools. The students in the suburban school also had significantly higher SERT quiz scores than either the rural or urban school. As such, one possibility is that the students in the suburban school may not have benefited from the training because they already had a good grasp on many of the SERT strategies (or other effective strategies) before training. In contrast, students in the rural school and to a lesser extent, students in the urban school, may have benefited because they were less familiar with the strategies before training.

These results have important implications for strategy instruction. First, our results highlight the importance of testing interventions in ecological and diverse settings. Simply testing the effectiveness of strategy in a laboratory setting does not necessary mean that it will scale up for use in classrooms. Moreover, success in a laboratory, or even a few classrooms, does not guarantee success in all types of high schools. Indeed, this research has led to some important cautions in providing strategy training to high school students. Second, researchers in the area a reading strategy instruction must be sensitive to the particular needs of the student. Students learn at different rates, so it is critical that instructors try to *tailor* their training so that students with diverse backgrounds can benefit. One weakness of the human-based SERT training used in this study was that not all of the students benefited from the training - only the students who mastered the SERT training in certain schools made significant gains. The effects of training may have been mitigated by the fact that it was given in a group, rather than individualized settings. Individual one-on-one training may produce stronger effects because the training can be adapted to the student's individual needs.

Our approach to solving this problem has been to automate the SERT training within a program called iSTART (McNamara, Levinstein, & Boonthum, 2004). iSTART is a computer program that uses automated agents to teach SERT. The program is adaptive and provides feedback on the quality of explanations based on the individual performance of the student. Preliminary results indicate that iSTART helps improve comprehension (McNamara, O'Reilly, Rowe, Boonthum, & Levinstein, in press). We speculate that iSTART is effective, in part, because it focuses on adaptive instruction that is tailored to students' needs.

In conclusion, our results provided evidence that, under certain circumstances, SERT can help students learn more from expository science text. However, care should be taken to ensure that students master the strategy before it should be evaluated. Finally our results underscore the importance of field testing reading interventions in diverse school settings.

Acknowledgements

References

Anderson, J.R. & Lebiere, C. (1998). *The atomic components of thought.* Mahwah, NJ: Erlbaum.

Ceci, S., & Bronfenbrenner, U. (1985). "Don't forget to take the cupcakes out of the oven": Prospective memory, strategic time-monitoring, and context. *Child Development, 56*, 152-164.

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

Finn, J., Gerber, S., Achilles, C., & Boyd-Zaharias J. (2001). The enduring effects of small class sizes. *Peabody Journal of Education, 67,* 75-84.

Fuchs, D., Fuchs, L.S., Mathes, P.G., & Simmons, D.C. (l997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal, 34,* 174-206.

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 trainer for active reading and thinking. *Behavioral Research Methods, Instruments, and Computers, 36*, 222-233.

McNamara, D.S., O'Reilly, T., Best, R., & Ozuru, Y. (in press). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research.*

McNamara, D. S., O'Reilly, T., Rowe, M., Boonthum, C. & Levinstein, I. (in press). iSTART: A Web-Based Tutor that Teaches Self-Explanation and Metacognitive Reading Strategies In D. S. McNamara (Ed.). *Reading*

Comprehension Strategies: Theory, Interventions, and Technologies. Erlbaum: Mahwah, NJ.

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 26th Annual Meeting of the Cognitive Science Society* (pp. 1053-1058). Mahwah, NJ: Erlbaum.

O'Reilly, T., Sinclair, G.P., & McNamara, D.S. (2004). Reading strategy training: Automated versus live. In K. Forbus, D. Gentner, T. Regier (Eds.), *Proceedings of the 26th annual meeting of the Cognitive Science Society* (pp. 1059-1064). Mahwah, NJ: Erlbaum.

Rivers, W. (2001) Autonomy at all costs: An ethnography of metacognitive self-assessment and self-management among experienced language learners. *Modern–Language Journal, 85*, 279-290.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64*, 479-530.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research, 66*, 181-221.

Shapiro, A.M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal, 41,* 159-189.

Stormont, M. (2005). Introduction to the special issue: Supporting learners with diverse needs in general education settings. *Psychology in the Schools, 42*(8), 759.

Van-Keer, H., & Verhaeghe, J. (2005). Effects of Explicit reading strategies instruction and peer tutoring on second and fifth graders' reading comprehension and self-efficacy perceptions. *Journal of Experimental Education, 73*, 291-329.