

Enhancing Learning of Expository Science Texts in a Remedial Reading Classroom via iSTART

Roger S. Taylor, Tenaha O'Reilly, Grant P. Sinclair, Danielle S. McNamara
Institute for Intelligent Systems, The University of Memphis, Department of Psychology

Abstract: Scientific literacy is a critical skill that requires the ability to deeply comprehend challenging scientific concepts. This research investigated the effectiveness of extended classroom practice with an automated reading strategy training program (iSTART) that is designed to enhance the comprehension of expository science text. Twelve 9th grade students, enrolled in a high school remedial reading program, were trained with the iSTART program. They were compared to a control group of twelve other 9th grade students, who were comparable in reading ability but who did not receive such training. There were no significant pretest-posttest differences between the groups on comprehension questions that could be answered from a single sentence within a passage. However, the iSTART training group performed significantly better than those in the control group on bridging inference questions, which require the generation of inferences between two or more sentences within a passage.

Introduction

The impact of science literacy on our economy and society in general is significant, and has been widely documented (e.g. AAAS, 1993; NRC, 1996). The recent report by the National Academy of Sciences (NAS, 2005) describes how American leadership in the areas of science and technology has been eroded, and lays out the serious economic problems that will result if this trend is not reversed. The committee rated as "most urgent" the key recommendation to "Increase America's talent pool by vastly improving K-12 mathematics and science education" (p. 4-2).

A primary mode of science instruction involves the reading of expository science texts. Given that the intended purpose of such texts is the introduction of new and unfamiliar concepts, it is not surprising that many students have difficulty reading such materials (Bowden, 1999; Snow, 2002). This problem is exacerbated because students do not typically use good reading strategies (Cox, 1997; Garner, 1990). Therefore, one promising approach is to improve students' ability to comprehend and learn from science texts through training in the use of more effective reading strategies. Empirical studies have demonstrated that interventions targeting the enhancement of reading strategies have been successful in improving reading comprehension (e.g. Pressley et al., 1992; Rosenshine & Meister, 1994; Rosenshine, Meister, & Chapman, 1996).

Recently, attention has been given to a reading intervention called Self-Explanation (Chi, De Leeuw, Chiu, & LaVancher, 1994). Self-explanation is a process in which a learner attempts to construct a more coherent mental representation of a poorly understood phenomenon, such as those in science texts. Previous research demonstrated impressive learning gains for students when they were prompted to self-explain (e.g., Chi et al., 1994). McNamara and colleagues have built upon this research by providing learners with scaffolded training in metacognitive reading strategies (McNamara, 2004; McNamara, Levinstein, & Boonthum, 2004). In the training program, called self-explanation reading training (SERT), students learned to use six metacognitive reading strategies: comprehension monitoring, paraphrasing, prediction, logic and common sense, elaboration, and bridging inferences. Overall, the results of SERT training indicate that it results in improved comprehension, particularly for low-knowledge readers (McNamara, 2004; O'Reilly, Best, & McNamara, 2004).

Interactive Strategy Training for Active Reading and Thinking: iSTART

While previous SERT results have been encouraging, there are two limitations. First, the high cost of human tutors makes widespread adoption of the program extremely difficult, particularly in economically distressed areas. Second, the structure of the SERT program does not allow for dynamic instruction that is tailored to the specific needs of the learner. Consequently, McNamara and colleagues developed an automatized version of SERT training called iSTART (Interactive Strategy Training for Active Reading and Thinking, see McNamara et al., 2004). Training with iSTART occurs in three phases and takes approximately 2 1/2 hours. The delivery of

information and feedback provided to the students is accomplished using animated pedagogical agents. The content of the feedback is determined based on the quality of the students' self-explanations, as assessed via a set of algorithms (for details, see McNamara et al., 2004).

The first phase of training is strategy *Introduction*. This phase includes definitions and examples of self-explanation and five reading strategies (i.e., comprehension monitoring, paraphrasing, prediction, bridging inferences, and elaboration). After each strategy is presented, students are asked to answer four multiple-choice questions and are then provided with immediate feedback by the program. The second phase of training is strategy *Demonstration*. In this phase, pedagogical agents model the use of the reading strategies while self-explaining a science text. Students are asked to identify the strategies used by the pedagogical agent during self-explanation, and again are provided with immediate feedback. The third phase of training is strategy *Practice*. In this phase, students read two short science passages and are asked to apply the newly learned strategies while typing self-explanations of the sentences in the texts. As they proceed through the text, several algorithms are employed to evaluate the quality (e.g., self-explanation length and the number and type of content words) of the generated self-explanations (McNamara et al., 2004). Based upon this analysis, scaffolded feedback is provided. For instance, a student providing an "impoverished" self-explanation that does not go beyond the content of the text's sentence is prompted by the agent to provide more information.

The iSTART system is adaptive and responsive to the user by providing feedback that is appropriate for specific level of the student. Investigations have revealed that using iSTART improves comprehension as compared to normal reading controls (O'Reilly, Sinclair, & McNamara, 2004a). Furthermore, research has shown the benefits of self-explanation training for low-knowledge readers, less-skilled readers, and for readers with little or no prior knowledge of reading strategies (Magliano et al., 2005; McNamara, 2004; McNamara, O'Reilly, Best, & Ozuru, in press; O'Reilly, Sinclair, & McNamara, 2004b). However, to date, there has been no study examining how self-explanation training affects the comprehension of less skilled readers over an *extended* period of post-training self-explanation practice.

Current Study

In the current study, we investigated the impact of iSTART in an ecologically valid, but highly challenging situation. That is to say, the administration at a local high school approached our research team for some assistance in helping with a remedial reading program. The administration had identified the 50 students exhibiting the lowest scores on an eighth grade reading comprehension proficiency exam, and offered them the opportunity to participate in a remedial reading class to enhance their reading comprehension skills. Twenty-five students were offered the class in the fall semester. Fifteen students accepted the offer.

Our goal was to integrate iSTART into the remedial reading class. At the beginning of the academic year we designed a curriculum that included iSTART training, followed by three weeks of extended practice using the iSTART reading strategies. We compared the iSTART-trained students to regular students on several comprehension measures at various points throughout the training/extended practice period to assess any beneficial effects of the iSTART training.

It is again emphasized that this study has high ecological validity because the intervention (i.e. pre-testing, iSTART training and extended practice, post-testing) was conducted by high-school teachers in their regular classes. In addition, all the texts used during the extended practice sessions were taken directly from the students' textbooks.

Method

Participants

The initial sample consisted of 93 9th grade students from a suburban Tennessee high school who were enrolled in a 9th grade life science course. Fifteen of these students, identified by the school as having reading comprehension difficulties based on an 8th grade reading exam, were enrolled in a special reading class designed to improve their reading comprehension. Three of the 15 reading class students did not complete the pretest reading ability measure and were, therefore, not included in these analyses. The remaining twelve students served as the *iSTART* group. Twelve students from the life science class were found who matched the *iSTART* students based on their pretest scores on a standardized measure of reading ability. This matched group of 12 students from the life

science class served as the *control* group.

Materials

Standardized Reading Ability. Reading ability was measured by the Gates-MacGinitie reading skill test for grades 7/9 (form K at pretest, form L at posttest). The Gates-MacGinitie is a 48 item, four-alternative, multiple-choice measure of reading skill designed to assess comprehension on several short text passages.

Prior Science Knowledge. Prior science knowledge was measured with a 20-item, four-alternative, multiple-choice test to assess students' science knowledge. The test covered several areas including biology, chemistry, earth science, research methods and mathematics. Questions were selected from high school science texts collected from several states (i.e., Colorado, Georgia, Kentucky, Tennessee, Virginia).

Pretest Passage Comprehension. The pretest passage was a 333 word passage on earthquakes which described the definition and causes of various types of earthquakes. The 25 sentence passage had a Flesch Reading Ease of 61.46 and a Flesch-Kincaid Grade Level of 8.0. An accompanying set of eight open-ended questions were developed for the passage. The answers to four of the questions could be found within a single sentence of the passage, and are referred to as *text-based* questions. The answers to the remaining four questions required the reader to combine information contained in two or more sentences of the passage, and are referred to as *bridging inference* questions. An example of a text-based question was "What is a subduction boundary?" while an example of a bridging inference question was "Describe how earthquakes occur according to the elastic-rebound theory."

Posttest Passage Comprehension. The posttest passage was a 284 word passage on red blood cells which described the structure and function of red blood cells in the human body. The 21 sentence passage had a Flesch Reading Ease of 59.06 and a Flesch-Kincaid Grade Level of 8.4. Eight open-ended questions (four text-based and four bridging inference questions), similar to those in the pretest, were developed based on the content of the posttest passage. An example of a text-based question was "How does sickle-cell disease get its name?" while an example of a bridging inference question was "Explain why blood plasma is a poor carrier of oxygen?"

Procedure

Prior to the experiment, the remedial reading class teacher attended a two day workshop that covered general principles of learning, memory, and reading comprehension, as well as a detailed lecture on iSTART and how to utilize it in the classroom.

The experiment consisted of four phases: pretest, training, extended practice, and posttest. The pretest and posttest phases were identical for both the iSTART and control conditions, while the iSTART training and extended practice were performed only by students in the iSTART condition.

During the pretest, students in the iSTART and control group were administered the pretest measures in the following order and time frame: prior science knowledge (10 minutes), Gates-MacGinitie reading measure (Form K, 15 minutes), and the passage on earthquakes along with the set of comprehension questions (15 minutes).

During the training, students in the iSTART group progressed through the three sections of the iSTART program over the course of three class periods on three consecutive days. Beginning the following week, participants in the iSTART condition entered the extended strategy practice phase. This phase took place over the course of three consecutive weeks. Each week the students self-explained two texts that were taken directly from the textbooks used in their regular courses. On day one, the students used the computer portion of iSTART practice to self-explain text one, with feedback being provided by the pedagogical agent in the same manner as during the practice phase of iSTART training. On the following day (day two), the teacher led the class in a self-explanation task of text 1, and provided feedback on performance. On day three, students used the practice section of iSTART to self-explain text 2, while on the following day (day four) the teacher led the practice for text 2. The time spent on practice each week on the two texts was approximately three hours, one and a half hours of iSTART practice and one and a half hours of teacher-guided practice. On the fifth day of the week, students were given quizzes on the material that had been covered during that week. The quizzes were also administered to the control subjects, who had studied the material in the classroom in the manner normally used by the life science teacher.

During the posttest, students in the iSTART and control groups were administered the posttest measures in the following order and time frame: Gates-MacGinitie reading measure (Form L, 15 minutes), and the passage on red blood cells along with the set of comprehension questions (15 minutes).

Results

Pretest Equivalence: Reading Comprehension and Prior Science Knowledge

The twelve students in the iSTART condition were matched with twelve students in the control condition based on the number of correct responses and the number of questions attempted on the standardized measure of reading comprehension (Gates-MacGinitie pretest scores). T-tests revealed no significant difference in pretest reading scores between the iSTART condition ($M = 9.75$, $SD = 4.29$) and the control condition ($M = 9.92$, $SD = 4.36$), $t(22) = 0.09$, $p = .93$ (2-tailed), and no significant difference in pretest prior science knowledge scores between the iSTART condition ($M = 0.35$, $SD = 0.09$) and the control condition ($M = 0.39$, $SD = 0.10$), $t(22) = 0.97$, $p = .34$ (2-tailed).

Text-Based Questions

Performance on the text-based questions was assessed by scoring the number of correct answers to the text-based questions as a proportion of the total number of text-based questions ($n = 4$). A mixed-model ANOVA was performed on the between-subjects factor of condition (iSTART, control) and the repeated-measures factor of time (pretest reading comprehension, posttest reading comprehension). Due to problems with the high-school computer network, the data for six students' pretest reading comprehension scores was incomplete. These observations were replaced with means based on the total sample¹. There was a significant increase in performance over time, $F(1, 22) = 10.6$, $p = .004$. There was no significant effect of condition, $F(1, 22) = 1.08$, $p = .311$. The interaction of time and condition was not significant, $F(1, 22) = 0.038$, $p = .847$.

Bridging Inference Questions

Performance on the bridging inference questions was assessed by scoring the number of correct answers to the bridging inference questions as a proportion of the total number of bridging inference questions ($n = 4$). A mixed-model ANOVA was performed on the between-subjects factor of condition (iSTART, control) and the repeated-measures factor of time (pretest reading comprehension, posttest reading comprehension). As noted earlier, missing data for six students' pretest reading comprehension scores for the pretest text were replaced with means based on the total sample². There was significant improvement in performance over time, $F(1, 22) = 58.15$, $p = .000$. There was a significant effect of condition, $F(1, 22) = 7.68$, $p = .011$. Most importantly, the interaction of time and condition was significant, $F(1, 22) = 7.65$, $p = .011$ (see Figure 1), and indicated that the students in the iSTART condition exhibited a greater increase in performance from pretest ($M = 2.00$, $SD = 2.09$) to posttest ($M = 35.42$, $SD = 14.92$) than that of the students in the control condition from pretest ($M = 2.08$, $SD = 7.22$) to posttest ($M = 17.71$, $SD = 14.56$). Additionally, the difference between the iSTART condition and the control condition at posttest was significant, $t(22) = 2.94$, $p = .004$.

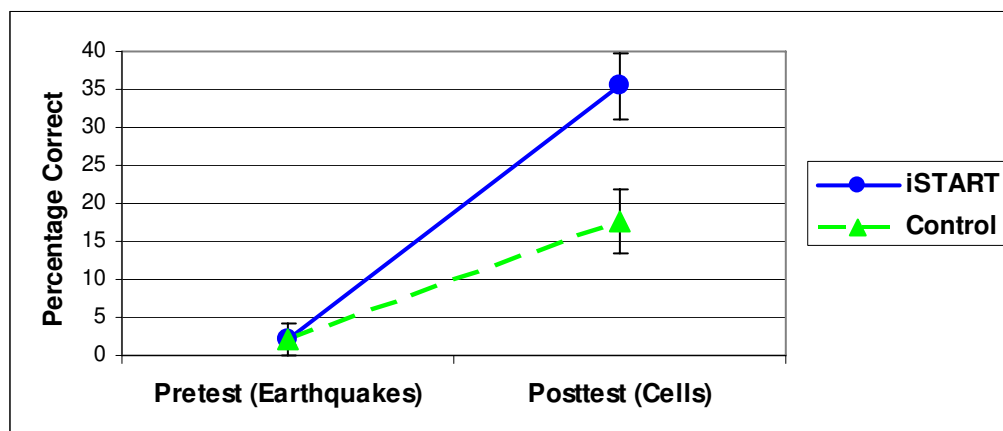


Figure 1. Percentage of correct bridging inference questions between conditions.

Discussion

The aim of this study was to examine the potential enhancement of student learning from expository science texts with extended classroom practice using the iSTART system. The key finding from this work was that this type of instructional intervention did in fact result in significantly greater performance on comprehension questions that required students to make inferences between different parts of a science text passage. This finding is especially important because this improved performance occurred on a more challenging aspect of comprehension.

During the course of the pretest phase, it was discovered that the school's computer system was not as reliable as had been believed, resulting in some data loss during that phase of the experiment. As a result of that discovery, the iSTART system has been reconfigured so that data collection can be managed independently of a school's computer network. Thus, this classroom experience helped us to identify a weakness in the software which, because we were able to address the problem, improves the viability of the iSTART program in school settings.

It should be noted that the passage for pretest comprehension assessment was on the physical science topic "earthquakes," while the passage for posttest assessment was on the biological science topic "red blood cells." A possible explanation for the higher overall scores on the posttest questions (red blood cells) is that the students in the study were enrolled in a life science class (a pre-biology course). Therefore, it is possible that the students were more familiar with the posttest materials. Nonetheless, it is clear that the iSTART trained students exhibited greater ability at answering the bridging inference questions after the training and practice phases of the experiment.

One of the main strengths of this study was its high ecological validity. Upon receiving a request to integrate the iSTART system into a local public high-school's remedial reading program, we were able to provide the teacher with the ability to conduct the training and practice in a normal classroom setting using texts from the student's textbooks. In the past, iSTART, and its feedback algorithms that are implemented during the practice phase, only included two training texts. This study indicated that iSTART practice can be readily extended to other texts.

A second encouraging factor is that, even with a limited number of students in a single remedial reading class, it is clear that the iSTART training had a significant facilitative impact on reading comprehension ability. While computer problems at pretest resulted in partial data loss, it is still clear that in comparison to the control group, the iSTART group performed significantly better on the posttest comprehension questions. Moreover, this result emerged on the more challenging bridging inference questions. In other words, the iSTART training helped these students, identified as needing remediation by the school system, to more deeply understand difficult expository science texts. We are currently performing more fine-grained analyses, including the hand-coding of the self-explanations, to more precisely determine the mechanism via which the training can enhance learning.

Previous research has suggested that iSTART training is effective because it improves the quality of students' elaborations (McNamara, O'Reilly, Best, and Ozuru, in press) and encourages students to use their general knowledge to construct deeper representations of the text (McNamara, 2004). The current results make sense with respect to these previous findings given that the more challenging bridging inference questions often require the use of elaborations and general knowledge.

With growing recognition of the national importance of science literacy, the ability to meaningfully read and learn from expository science texts is a critical educational issue. The results presented here demonstrate the value of using a powerful educational tool such as the iSTART instructional system.

Endnotes

- (1) The results were the same when the missing values were replaced with the matched students' scores.
- (2) The results were the same when the missing values were replaced with the matched students' scores.

References

American Association for the Advancement of Science (AAAS), Project 2061. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

- Bowen, B. A. (1999). Four puzzles in adult literacy: Reflections on the national adult literacy survey. *Journal of Adolescent & Adult Literacy*, 42, 314-323.
- Chi, M. T. H., De Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18, 439-477.
- Cox, B. D. (1997). The rediscovery of the active learner in adaptive contexts: A developmental-historical analysis of transfer of training. *Educational Psychologist*, 32, 41-55.
- Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. *Review of Educational Psychology*, 60, 517-529.
- Magliano, J. P., Todaro, S., Millis, K. K., Wiemer-Hastings, K., Kim, H. J., & McNamara, D. S. (2005). Changes in reading strategies as a function of reading training: A comparison of live and computerized training. *Journal of Educational Computing Research*, 32, 185-208.
- McNamara, D. S. (2004). SERT: Self-Explanation Reading Training. *Discourse Processes*, 38, 1-30.
- McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iSTART: Interactive strategy training for active reading and thinking. *Behavior Research Methods, Instruments, & Computers*, 36, 222-233.
- McNamara, D. S., O'Reilly, T., Best, R., & Ozuru, Y. (in press). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research*.
- National Academies of Sciences (NAS). (2005). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC: National Academy Press.
- National Research Council (NRC). (1996). *National science education standards*. Washington, DC: National Academy Press.
- 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.P., Sinclair, G.P., & McNamara, D.S. (2004a). 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.
- O'Reilly, T.P., Sinclair, G.P., & McNamara, D.S. (2004b). iSTART: A web-based reading strategy intervention that improves students' science comprehension. In Kinshuk, D. G. Sampson, & P. Isaias (Eds.), *Proceedings of the IADIS International Conference Cognition and Exploratory Learning in Digital Age: CELDA 2004* (pp. 173-180). Lisbon, Portugal: IADIS Press.
- Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist*, 27, 91-109.
- 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.
- Snow, C. (2002). *Reading for understanding: Toward an R & D program in reading comprehension*. Santa Monica, CA: RAND.