ED-MEDIA 2003

World Conference on Educational Multimedia, Hypermedia & Telecommunications

Edited by
David Lassner
Carmel McNaught

Proceedings of ED-MEDIA 2003

WORLD CONFERENCE ON EDUCATIONAL MULTIMEDIA, HYPERMEDIA & TELECOMMUNICATIONS

June 23-28, 2003; Honolulu, Hawaii, USA

AACE
Association for the Advancement of Computing in Education
A CONFERENCE OF AACE
http://www.aace.org

Web-Based Intervention for Higher-Order Reading Skills

I. B. Levinstein (ibl@cs.odu.edu), Dept. of Computer Science, Old Dominion University, Norfolk, Virginia, USA

D. S. McNamara (d.mcnamara@mail.psyc.memphis.edu), Dept. of Psychology, The University of Memphis, Memphis, Tennessee, USA

C. Boonthum (cboont@cs.odu.edu), Dept. of Computer Science, Old Dominion University, Norfolk, Virginia, USA

S. P. Pillarisetti (s.pillarisetti@mail.psyc.memphis.edu), Dept. of Psychology, The University of Memphis, Memphis, Tennessee, USA

K. Yadavalli (kyadaval@cs.odu.edu), Dept. of Computer Science, Old Dominion University, Norfolk, Virginia, USA

Abstract: A promising metacognitive reading strategy intervention is being evaluated in several high schools as a means of improving students' comprehension of science texts. Students are trained to use a variety of strategies in explaining texts to themselves when they do not understand a sentence or paragraph. To make the training more widely available, a web-based trainer has been developed. Transforming the training from human- to computer-based resulted in a highly interactive trainer that adapts its methods to the performance of the students. The iSTART trainer introduces the strategies in a classroom setting with interaction among three animated characters, a teacher and two student characters, and the human trainee. Then the trainee identifies the strategies in the explanations of a student character guided by a teacher character. Finally, the trainee practices explaining under the guidance of a teacher character. We describe this system and discuss how appropriate feedback is generated.

I. Introduction

Many high-school students in the United States have difficulties understanding scientific material in their textbooks even when they have mastered the basic decoding and syntactic skills of reading. One source of this problem is that they fail to use higher order reading strategies when they do not understand the text. Self-explanation is a strategy that improves comprehension of difficult text (Chi et al., 1994). McNamara (2002) has developed a reading strategy intervention called Self Explanation Reading Training (SERT), delivered by human trainers in a classroom setting that has shown promising results with college freshmen and high-school students. During SERT, students are trained to use several reading strategies while self-explaining difficult text. These strategies improve the quality of their self-explanations, which in turn improves comprehension. We are currently exploring the use of this intervention and others in a project involving several hundred high school students from diverse communities.

The cost of delivering these reading strategy interventions using human trainers is relatively high and can be provided to a limited number of students. In order to make the intervention available to a large number of students with minimal cost, we are developing a web-based version of the intervention called Interactive Strategy Trainer for Active Reading and Thinking (iSTART) that delivers the training in a one-to-one fashion. This article discusses the transformation of a classroom-oriented training program into an electronic one-on-

one trainer. In the sections that follow we describe the human-led training and the constraints on the iSTART system. This is followed by a description of the iSTART system, its pedagogy and its means of evaluating trainees' explanations.

II. The SERT methodology

The aim of SERT (Self Explanation Reading Training) is to provide readers with strategies they can employ when they fail to understand a text. The target population of SERT consists of high school and college students who have adequate decoding and syntactic skills but who still have difficulties with complex texts. SERT is one of several metacognitive reading methodologies, so called because they encourage readers to direct their own thinking in ways that will facilitate reading. Other metacognitive training programs emphasize acquiring a global overview of the text's concepts by pre-reading (Previewing, Richardson & Morgan, 2000) or periodically monitoring one's comprehension of the text (INSERT, Vaughan & Estes, 1986). We are currently researching the application of metacognitive training in a variety of high school environments.

SERT trains readers to use a variety of active reading strategies to self-explain difficult text more effectively. The first and indispensable strategy is *comprehension monitoring*, the habit of being aware of how well one is understanding the text. Comprehension monitoring enables the reader to recognize a failure of understanding and it is this recognition that triggers the use of additional reading strategies. SERT training recommends that the first of these should be rendering the sentence in the reader's own words or *paraphrasing*. Although paraphrasing by itself does not constitute an explanation it helps to ensure that the reader understands the sentence's grammar and vocabulary. Paraphrasing may be enough to trigger comprehension; it may also alert the reader that some words are unfamiliar and that their meanings need to be deduced from context, discovered from later material in the text, or looked up. Students are trained in SERT to go beyond sentence-focused paraphrasing by invoking knowledge-building strategies that link the content of the sentence to other material. For example, they are encouraged to use a strategy called *bridging* to tie the current sentence to the material previously covered in the text. In addition, they may use *prediction* to anticipate the content of the rest of the text, either by guessing what is coming next or by reminding themselves to watch out for some particular item that will aid comprehension. Finally, they may associate the current sentence with related knowledge, gained from sources outside the text, using a strategy called *elaboration*.

The human-based training for the SERT intervention, apart from pre- and post-testing and possible refresher sessions, takes about two hours of classroom time in three phases: introduction, demonstration, and practice. In our experiments, science teachers in participating high schools donate the time from their regular class periods. The introduction is a lecture that motivates the training and describes the strategies. The trainees are told that the training may improve their ability to understand science texts and therefore to do better in school. The lecture defines and provides examples of each of the strategies. In the demonstration section that follows the introduction, the trainees are given a text and shown a video of a student reading and self-explaining it aloud. The video is interrupted at several points and the trainees are asked to identify the strategies that the student in the video has been using. The student in the video models the desired self-explanation behavior and the discussion of strategies clarifies points that had been presented more abstractly during the lecture. The demonstration gives the trainees concrete instances of the behaviors that the lecture introduced. The discussion aids them in understanding and recognizing those strategies by encouraging them to identify the strategies in the explanations demonstrated in the video. In the third section, practice, the students pair up under the supervision of the trainer and self-explain passages to each other.

A series of experiments, conducted with college freshmen and high school students in three different school districts, has yielded statistical evidence that the SERT training improves student performance in science courses to a significant degree and more consistently than either the control groups or those trained in other metacognitive approaches.

III. The iSTART trainer: Design considerations

The iSTART trainer (Interactive Strategy Training for Active Reading and Thinking) was created to provide SERT training in a more scalable and cost effective manner than is possible using human trainers and to provide an intervention that could be used outside of normal class time. Some of the constraints on the project were the same as on the human-based training. The training had to be designed as a short intervention and readily adaptable to a variety of science disciplines. The reason for both is that high-school science teachers have a great deal of material to cover during the year. It is difficult to persuade them to adopt a training program that takes class periods away from lessons focused on course material and they are more willing to accept one that involves material from the subject they are teaching. In addition, we needed to be able to make frequent revisions to the trainer and to collect detailed data on the interactions of the trainees with the software. For this reason, iSTART was designed to be accessed via an Internet browser from a central server that could record all student interaction in a database.

In designing iSTART we wanted to maintain the basic character of the human-based training, build on its strengths, and adapt it to the web environment. The basic character of the original training lies in the three phases: 1) an introductory explanation which presents definitions and examples, 2) a demonstration of the techniques in action coupled with the trainees' use of those concepts in an analytical fashion, and 3) an opportunity to practice the techniques under guidance. A major strength of this training is the variety of ways in which the trainees interact with the SERT strategies. They receive them as a passive audience for definition and example; they see them in use by someone like themselves; they analyze those uses in a group discussion; they observe and prompt another trainee who is using the techniques; and they practice the techniques themselves. We wanted to preserve the phases and variety of the human-based training when we developed iSTART for the web environment. In that environment, the computerized trainer would deal with one trainee at a time, rather than with a classroom. While the trainers in the classroom gained authority from the regular classroom teacher and from their position in front of the class, iSTART would have no such advantage. In addition, trainees with computer experience have expectations far different from what they have in the classroom. Finally, the human trainers and fellow trainees come equipped with enough intelligence to respond appropriately to the attempts at self-explanation made during the practice sessions, an ability which takes a good deal of programming in a computer application.

The web environment afforded many opportunities for improving the human-led classroom training. Because it deals with trainees one-at-a-time, the training can be self-paced and adaptive to the trainee. More importantly, the training can engage the trainee in more interactive dialog and require the trainee to be a more active learner. In developing the trainer we faced two major problems: a) how to maintain the trainees' interest and involvement with the training process while still being effective (the pedagogical problem), and b) how to determine the quality of the trainees' explanations so that the trainer could respond appropriately (the evaluation problem). The pedagogical problem required the replacement of passive aspects of the human-based training with interactivity appropriate to a computer-based setting. The evaluation problem was complicated by the fact that we had to recognize a trainee's typed response as constituting one strategy or another or as exhibiting a quality rather than simply determining whether an answer was correct or incorrect.

IV. iSTART trainer: Animated pedagogy

iSTART follows the order of presentation used in the human-based SERT training but the training approach is adapted to the computer-based setting. The human-based training is delivered in a classroom and uses group oriented lectures and demonstrations. Ordinary classroom discipline and social pressure suffice to keep trainees' attention during these presentations. In a computer-based environment, more than a talking head is required to maintain that attention. We chose to employ animated talking characters to deliver the information and guide the trainee. The characters we designed have a variety of gestures and speak using a text-to-speech synthesizer. As they speak, their words appear in a balloon above their heads. The only drawback is that the characters and text-to-speech engine must be installed on the trainees' computers. The characters' behavior is controlled by scripts embedded in web pages, so it is easy to modify the trainer without necessitating any additional installation. All the modifications are made to pages on the server.

The human-based training consists of a lecture that motivates SERT and introduces the trainee to its basic concepts and techniques. In the iSTART introductory module, we replace the human-delivered lecture with a classroom discussion between an instructor character, two student characters, and the trainee (see Figure 1). This format provides more variety in the ways in which the trainee interacts with the concepts and strategies than does the human training. These interactions take the following forms:

1. The instructor character presents material as in a lecture with both definitions and examples.
2. The instructor questions the student characters.
3. The instructor questions the trainee.
4. The students banter among themselves.
5. The students ask the instructor for examples or clarifications.
6. The students ask the trainee for help.

The introduction is broken down into a number of sections and the trainee has to click on a button to move from one to the next. In addition, the trainee is given brief quizzes at several points during the introduction and advised to continue to the next section or to review the last depending on the results. By all of these means, the trainee is induced to be an active participant in the presentation. There is little opportunity for the trainee's attention to wander. And, should the trainee desire, he/she can easily return to an earlier part of the introduction for review.

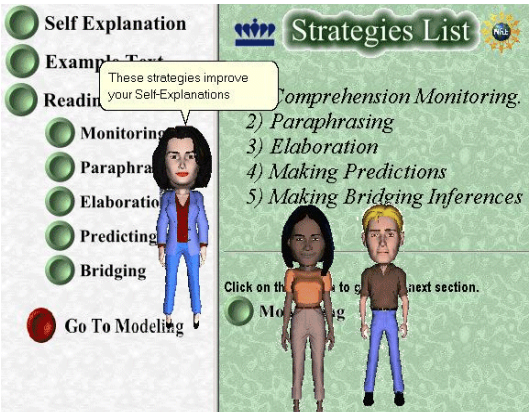

Figure 1: Screen Shot from Introduction Module

The theme of interactivity is carried into the demonstration module of iSTART (see Figure 2). In the human-based training, the training group is shown a video of a student self-explaining a text. The video is interrupted several times for a discussion concerning which strategies were used in an explanation. The same basic approach is taken in iSTART. The text is read and explained sentence-by-sentence by a student character under the supervision of a teacher character. Sometimes the teacher prompts the student character to revise the explanation by:

1. requiring the student to add to an explanation that is too short,
2. requiring the student to revise an explanation that is too similar to the sentence being explained, or
3. encouraging the student to expand the explanation if it seems unrelated to the target sentence.

The trainee may encounter the same prompts in the later practice session. After the student character has produced a satisfactory explanation, the teacher character switches its attention to the trainee. Generally the teacher asks the trainee to decide which strategies were used in the student character's explanation. At times the teacher will ask follow-up questions that focus the trainee's attention on the explanation and the target sentence. The teacher will ask the trainee to identify with a mouse-click where the student character used a certain strategy. If the trainee correctly identifies an occurrence of bridging, the teacher may ask the trainee to click on the sentence to which the explanation is linked. If the explanation uses several strategies, the teacher may ask the trainee to identify another strategy or prompt that a certain strategy was used and ask the trainee to locate it. Whereas the human-based training only provides an interactive discussion of three or four explanations during a training session, iSTART interacts with the trainee after every explanation.

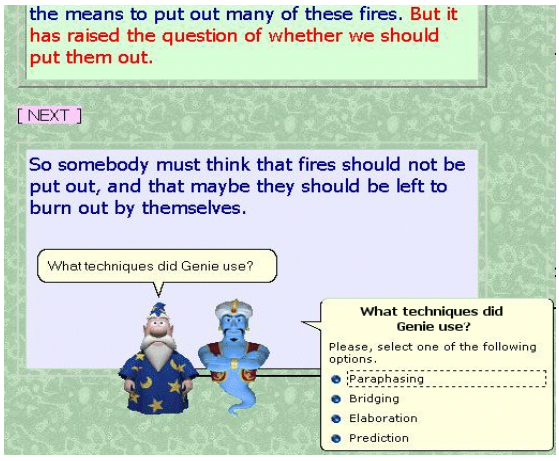

Figure 2: Screen Shot from Demonstration Module

The demonstration module dynamically adapts to the success of the trainee by varying the style of questioning that the teacher uses. The trainee's success is measured by the number of correct answers provided in the last few interactions. If the success level is low, more supportive styles of questioning are used. If it is high, less follow-up questioning is required. The range of questioning styles, from most to least supportive, includes:

1. Tell the trainee that a certain strategy was used, remind the trainee of the definition of that strategy, and ask the trainee to identify where it is used by clicking on the text of the explanation.
2. Like number 1, but omit the definition.
3. Present an explanation as a series of sub-explanations and, for each sub-explanation, ask the trainee to identify the strategy used by selecting it from a list. Ask follow-up questions in detail.
4. Like number 3, but without breaking down the explanation into sub-explanations. Usually ask the trainee to find two different strategies. Ask follow-up questions in detail.
5. Like number 4, but with fewer follow-up questions.

Trainees usually begin the demonstration module being questioned in multiple-choice (i.e., number 4) style. This is a moderately difficult task since one explanation may use several strategies. In this mode the trainees have to discriminate among strategies while they discriminate among sections of the explanation. If they identify a strategy used in the text, they may be asked one or two follow-up questions such as "Click on the part of the explanation where that technique was used" or "Click on the part of the text to which this explanation linked." If they are not successful, the trainer switches to less demanding modes in which explanations are broken down into smaller units or trainees are told that a particular strategy is used and then asked to locate its position in the explanation. Trainees who are successful with the multiple-choice questioning with respect to longer explanations are presumed to have a good understanding of the techniques. Once they reach this level, it is not necessary to verify their understanding by persisting in follow-up questions, so fewer and fewer follow-ups are used as long as the trainees continue successfully.

V. Evaluation of trainee responses.

In the demonstration section, the trainees have analyzed the self-explanations of a surrogate trainee (i.e., the student character), who has used the full range of reading strategies. Thus the demonstration section provides the trainee with a mental model of self-explanation that they can use when they are asked to explain a text in the practice module. In the human-based training, trainees are paired up and asked to coach each other as they practice explanations of a selected passage of a science text. In iSTART, the teacher character interacts with the trainee. The interaction follows the pattern that the trainee has observed in the demonstration section. During the practice section, the trainee self-explains sentences from texts, while attempting to use the reading strategies learned during the introduction and demonstration sections.

The programming challenge faced during this phase is to provide appropriate feedback to the trainee concerning the quality of the self-explanations. We have approached this 'evaluation' challenge in three ways. First, we evaluate the trainee's response to verify whether the explanation is a) sufficiently long, b) sufficiently different from, and c) relevant to the sentence and/or topic. If any of these three criteria is not met, then the trainee is asked to modify the explanation by adding more information or more details (to add relevance). Second, the trainee's explanation is evaluated in terms of its quality to guide the general feedback provided by the teacher agent. For example, if the self-explanation was relevant, and long enough, but categorized as a simple paraphrase, the teacher agent would give reserved feedback (e.g., "OK") or ask the student to try harder for the next sentence. On the other hand, if the self-explanation were long, relevant, and included many words that were not taken directly from the sentence, the teacher agent would likely respond very positively to the student (e.g., "Excellent job!"). The third way that we have achieved the goal of evaluating trainee's responses uses a more direct strategy: we simply ask the trainees what strategy they have used. This direct approach achieves two goals. First, we are putting the trainee in the position of metacognitive self-evaluation. This forces the trainee to think about the self-explanation more objectively and to think more about the different types of strategies that might have been used. Second, by using this approach we will be able to retrospectively assess the ability of both the trainees and our system to categorize self-explanations (i.e., by comparing these assessments to experimenter ratings).

Throughout the interaction, the module tracks the variety of strategies employed. Selected sentences have been identified as particularly appropriate for certain strategies. When they come to one of these sentences, the trainee may be encouraged to use that particular strategy if it has not been used during the practice.

The aim of the practice module is that the trainee employs the strategies while self-explaining a difficult text. Having the trainee identify the strategies used in his/her own self-explanations is simply a means to that end, but too much interaction slows the trainee's progress from one sentence to another. Therefore, the module gradually adopts a less intrusive posture than that of the demonstration section during which every sentence was explained and every explanation analyzed. At first, iSTART asks them to identify the strategies used in their self-explanations and continues with follow-up questions. If the practice proceeds successfully, the teacher character becomes less and less intrusive, perhaps just prompting the trainee to go on to the next sentence or occasionally asking the trainee to expand an explanation that seems too short or irrelevant. In this way the trainee can get on with the main task of this section, which is of course practice.

VI. Concluding Discussion

A primary interest of ours was the success of our algorithm for evaluating student responses, which first flagged responses as being too short, too similar to the sentence being explained, or apparently unrelated to that sentence and then, if no flags were raised, judged the quality of the sentence on a scale of 1 to 3. The student responses were collected during the experiment and later evaluated by a trained graduate student. It appears that we can claim only limited success for our algorithms at present. Of the three reasons for flagging, excessive similarity was by far the most frequent. Of those responses iSTART judged to be too similar, 50% were also found too similar by the human evaluator; and of those judged too similar by the human, iSTART agreed 50% of the time. For unflagged responses, the correlation between the human and program evaluations was 0.5. While this is a high correlation in some fields of research, it is not nearly reliable enough to use a basis for guiding trainees. We are actively experimenting with alternatives.

Trainees generally responded favorably to iSTART. We surveyed the subjects after two experiments with iSTART, one with a group of college students, mostly freshmen, taking an introductory biology course, and one with a group of middle school children, who were participating in a summer enrichment program for above-average students. Most of each group (college, 71%; middle school, 51%) recognized that they had previously used one or more of the reading strategies. Both groups rated iSTART highly (4 on a scale of 5) on how much they had learned. College students generally found the practice module the most useful (55%) while middle schoolers were split between practice (45%) and demonstration (40%). Their written comments suggest that they enjoyed working with iSTART even as they made suggestions for its improvement.

The histories of trainee interaction that we have collected during these and other experiments guide us in the development of the system. For example, in these histories we discovered that some trainees repeatedly tried to paraphrase sentences that were apparently too difficult for them to understand. They could not put the sentence into their own words so they repeated it with very minor changes. Rather than simply asking the trainees again and again to extend their explanations as iSTART had been doing, it will incorporate a routine that comes to their assistance encouraging them to use other strategies in that situation. Moreover, we can expect improved evaluation algorithms because we can test them on a corpus of hundreds of explanations.

VI. References

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

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,* 61-83

Graesser, A. C., Franklin, S. P., Wiemer-Hastings, P., & the Tutoring Research Group (1998). Simulating smooth tutorial dialogue with pedagogical value. In *Proceedings of the 11th International Florida Artificial Intelligence Research Symposium Conference*, pp. 163-167. AAAI Press.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes, 25*, 259-284.

McNamara, D. S. (2002), SERT: Self-Explanation Reading Training, *Journal of Educational Psychology* (submitted).

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.

National Assessment of Educational Progress (1998). *Education Commission of the States*, Denver, CO.

Richardson & Morgan, R. (2000), *Reading to Learn in the Content Areas*, CA: Wadsworth Publishing Co

Vaughan, J. & Estes, T. (1986). *Reading and reasoning beyond the primary grades.* Allyn and Bacon.

Voce, Julie & Hamel Marie-Josée (2001) Towards a Compromise Between Talking Heads & Interface Agents: A Web-based "Mentor" for Computer Assisted Language Learning (CALL), *World Conference On Educational Multimedia, Hypermedia & Telecommunications, 2001*, Association for the Advancement of Computing in Education, Tampere, Finland, 1964-1965

Zellweger, P. & Mackinlay, J.,(2001) The Fluid Reading Primer: Animated Decoding Support for Emergent Readers, *World Conference On Educational Multimedia, Hypermedia & Telecommunications, 2001*, Association for the Advancement of Computing in Education, Tampere, Finland, 2124-2129.

VII. Acknowledgement

This work was supported by National Science Foundation grant REC-0089271.