Journal of Educational Psychology
2000, Vol. 92, No. 4, 724–733

0022-0663/00/$5.00 DOI: 10.1037//0022-0663.92.4.724

Engaging Students in Active Learning: The Case for Personalized Multimedia Messages

Roxana Moreno and Richard E. Mayer
University of California, Santa Barbara

The authors tested the hypothesis that personalized messages in a multimedia science lesson can promote deep learning by actively engaging students in the elaboration of the materials and reducing processing load. Students received a multimedia explanation of lightning formation (Experiments 1 and 2) or played an agent-based computer game about environmental science (Experiments 3, 4, and 5). Instructional messages were presented in either a personalized style, where students received spoken or written explanations in the 1st- and 2nd-person points of view, or a neutral style, where students received spoken or written explanations in the 3rd-person point of view. Personalized rather than neutral messages produced better problem-solving transfer performance across all experiments and better retention performance on the computer game. The theoretical and educational implications of the findings are discussed.

Classic research in psychology has shown that people react differently to situations that involve personal reference. This is the case of the well-known *cocktail party effect,* in which a person who is attending to one conversation is able to detect his or her own name in a separate conversation that is taking place simultaneously in the same room (Cherry, 1953). Another robust phenomenon is the *self-referential effect,* in which retention is facilitated by having people process information by relating it to aspects of themselves (Rogers, Kuiper, & Kirker, 1977).

As with other factors in human learning and memory, the question arises as to whether personal reference effects can be extended and used in the area of multimedia learning. For example, it is possible to teach a computer-based science lesson with higher or lower levels of self-referencing by varying the communication style used in the explanations (Turco, 1996). In the studies reported in the literature, high self-referencing has been created by directly addressing students and encouraging them to believe that they were active participants in the lesson. For example, in a computer game aimed at teaching environmental science, a high level of self-referencing is present when an on-screen agent speaks to the learner as follows:

> This is a very rainy environment, and the leaves of your plant have to be flexible so they're not damaged by the rainfall. What really matters for the rain is your choice between thick leaves and thin leaves. Which do you think would be more flexible?

On the other hand, low self-referencing has been created by giving the same message without directly addressing students or encouraging them to believe that they were participants in the lesson, such as in the following explanation: "In very rainy environments plant leaves have to be flexible so that they are not damaged by the rainfall. What really matters for the rain is the choice between thick leaves and thin leaves."

Does the use of personalized messages improve students' understanding of a multimedia science lesson? To answer this question, in the present studies we tested the hypothesis that personalized explanations can promote deep understanding of a computer-based lesson by actively engaging students in the elaboration of the materials. Our hypothesis is based on two assumptions. First, we assume that the use of self-referential language promotes the elaboration of the instructional materials (Symons & Johnson, 1997). Second, we assume that participants use less cognitive effort to process verbal information when it is presented in a familiar style (i.e., normal conversation) rather than an unfamiliar style (i.e., monologue) of communication (Mayer, 1984; Spiro, 1977).

To test our hypothesis, we compared the learning outcomes of college students who learned a science lesson from personalized conversations with those of students who learned the same materials from neutral monologues. Two multimedia learning scenarios were used: a multimedia explanation about the process of lightning formation and an agent-based science computer game.

Roxana Moreno and Richard E. Mayer, Department of Psychology, University of California, Santa Barbara.

Roxana Moreno is now at the Educational Psychology Program, University of New Mexico.

This research was supported by a grant entitled "Learning and Intelligent Systems: Animated Pedagogical Agents for Constructivist Learning Environments" from the National Science Foundation. James C. Lester and his colleagues (Lester, Stone, & Stelling, 1999) are the original authors of the Design-a-Plant computer program. All modified versions of the original Design-a-Plant computer program were developed by Roxana Moreno. Jeff Canatta, Josh Kashinsky, and James Rolfe assisted by recording the voice of the pedagogical agents used in Experiments 1, 2, and 3, and 4 and 5, respectively.

Correspondence concerning this article should be addressed to Roxana Moreno, Educational Psychology Program, Simpson Hall 123, University of New Mexico, Albuquerque, New Mexico 87131 or to Richard E. Mayer, Department of Psychology, University of California, Santa Barbara, California 93106. Electronic mail may be sent to moreno@unm.edu or mayer@psych.ucsb.edu.

To what extent can the self-referencing of instructional materials help students' learning? The main focus of our study was to examine self-reference effects on the comprehension of a multimedia explanation. However, because most of the past research on self-reference has focused on its mnemonic effects (Symons & Johnson, 1997), we used two measures of learning: retention—in which we assessed memory for the basic factual information that was presented—and problem-solving transfer—in which we asked students to solve new problems on the basis of the principles learned in the respective computer program.

From Self-Referencing to Personalizing Multimedia Messages

Our first assumption is that self-referential language promotes the elaboration of the instructional materials (Symons & Johnson, 1997). The self-referential effect has enjoyed a long history in psychology. Although various accounts of this effect differ in important ways from one another, they do agree on one basic point: People remember information better when it is encoded with respect to themselves rather than with respect to other frames of reference (Rogers et al., 1977). In a meta-analysis of self-reference effects in memory, Symons and Johnson (1997) examined studies that used the depth-of-processing incidental-learning paradigm (Craik & Tulving, 1975), in which an encoding question is posed and a stimulus word is presented for the participant's judgment: The self-reference effect on memory is based on a very efficient mechanism to process material that is very familiar to oneself, "material that is often used, well organized, and exceptionally well elaborated" (Symons & Johnson, 1997, p. 392).

The facilitative effects of self-reference on retention have been observed across different types of stimuli (e.g., faces; Mueller, Bailis, & Goldstein, 1979), tasks (e.g., reading prose with self-referencing processing instructions; Reeder, McCormick, & Esselman, 1987), and contexts (e.g., persuasion; Burnkrant & Unnava, 1989; Hoban & van Ormer, 1970). In addition to its memorial effects, self-referencing improves the comprehension of mathematical word problems by elementary students (Anand & Ross, 1987; d'Ailly, Simpson, & MacKinnon, 1997; Davis-Dorsey, Ross, & Morrison, 1991; Ross, McCormick, & Krisak, 1986). Personalizing the context improves learning by helping learners interpret and interrelate important information in the familiar versus abstract problem statements (Mayer, 1984; Ross, 1983).

Personalized Messages as Minimizers of Effort

Our second assumption is that personalized messages are more consistent with our schemas for communicating in normal conversations and therefore require less cognitive effort to process. For example, in a study on note taking and passage style, students who read a text passage written in a low-formality style recalled significantly more than those who read the same passage written in a high-formality style (Bretzing & Kulhavy, 1981). The mnemonic superiority resulting from relating the material to the self was interpreted as originating in the easier processing required to update current schemata from a more familiar style of prose (Spiro, 1977).

In addition, the idea that familiarity may explain why some language styles are easier to comprehend and remember than others has been reported by a number of studies of narrative comprehension (Graesser, Golding, & Long, 1998). Narrative discourse is easy to comprehend and to remember, compared to other discourse genres, such as definition, description, and exposition (Freedle & Hale, 1979; Graesser, 1981; Spiro & Taylor, 1987). Moreover, in the area of communications, Reeves and Nass (1996) provided convincing evidence that people have a natural predisposition to apply the same dynamics from human–human interactions to human–computer interactions. Therefore, providing personalized messages in media communication seems more likely to ease the processing of the message by being more consistent with the social rules and schemas of normal conversations.

Hypothesis

The design of multimedia explanations can be driven by the designer's conception of the nature of teaching, which can range from an interactive hypothesis to a transmission hypothesis. We propose an interactive hypothesis of teaching according to which the depth of processing of an instructional message depends, among other factors, on the level of mental interaction promoted in the student by the language style used by the teacher. According to the interactive hypothesis, providing explanations by means of self-referenced dialogues rather than depersonalized generic monologues will encourage the learner to actively search for meaning (Anderson & Pearson, 1984; Doctorow, Wittrock, & Marks, 1978). Consequently, this approach predicts that students who learn a computer-based lesson by means of a personalized message will remember more of the factual information and solve problems better than students who learn by means of neutral messages.

On the other hand, according to the transmission hypothesis, human communication involves three processes: first, encoding an idea into a signal by a sender; second, the transmission of the signal to the receiver; and third, the decoding of the signal by the receiver (Reynolds, 1998). When the transmission hypothesis is applied to instructional messages, the role of the teacher is to present information, and the role of the learner is to receive information. Accordingly, this approach predicts that students who learn a multimedia science lesson by means of a personalized message or a neutral message will not show significant differences in their learning performance. The rationale for this prediction is that both programs contain identical visual information and equivalent verbal materials, which differ only in style.

Experiment 1

Many human–computer interactions involve a one-on-one learning scenario, that is, an instructional situation involving one human interacting with one computer. An example of a one-on-one learning scenario consists of a single student looking up an entry in a multimedia encyclopedia (such as an entry on the process of lightning formation). The purpose of Experiment 1 was to determine whether the self-reference learning effects found in the past (Symons & Johnson, 1997) would extend to a simple kind of multimedia-learning task. In Experiment 1 we compared students' learning from a short multimedia explanation of how lightning storms develop in which an animation was accompanied by either neutral or self-referenced speech.

Method

Participants and design. The participants were 34 college students recruited from the psychology subject pool at the University of California, Santa Barbara. There were 17 participants in the personalized-speech group and 17 participants in the neutral-speech group. All participants indicated a low level of knowledge of meteorology.

Materials and apparatus. The paper materials consisted of a participant questionnaire, retention test, and four problem-solving transfer test sheets, similar to those we have used previously (Mayer & Moreno, 1998). The participant questionnaire solicited the participant's age, gender, and SAT scores; in addition, to measure the participant's knowledge of meteorology, it contained the following two questions: (a) "Please put a check mark indicating your knowledge of meteorology (weather)," followed by five blanks ranging from *very little* (scored as 1 point) to *very much* (scored as 5 points), and (b)

> Please place a check mark next to the items that apply to you: ____ I regularly read the weather maps in a newspaper, ____ I know what a cold front is, ____ I can distinguish between cumulous and nimbus clouds, ____ I know what a low pressure system is, ____ I can explain what makes the wind blow, ____ I know what this symbol means [symbol for cold front], ____ I know what this symbol means [symbol for warm front].

The retention test had the following sentence at the top of the page: "Please write down an explanation of how lightning works." Each problem-solving transfer test sheet contained one of the following questions at the top of the page: "What could be done to decrease the intensity of a lightning storm?," "What does air temperature have to do with lightning?," "Suppose you see clouds in the sky, but no lightning. Why not?," and "What causes lightning?" The following instruction was typed at the bottom of each test sheet: "Please keep working until you are told to stop."

The computerized materials consisted of two computer programs that presented a 140-s animation depicting the steps in the formation of lightning along with corresponding narration describing the steps. The neutral-speech (N) version involved speech using only the third person (as presented in the body of Appendix A), and the personal speech (P) version involved speech in the first and second person as well as comments directed at the learner (with additions indicated in the brackets and deletions indicated by underlining in Appendix A). The programs were produced using Director 4.0.4 and SoundEdit 16, version 2 (Macromedia, 1995), and the neutral-speech version was identical to the animation and narration (AN) version used in Mayer and Moreno (1998).

The apparatus consisted of five Macintosh IIci computer systems with 14-in. (35.6-cm) color monitors and Koss headphones.

Procedure. Participants were tested in groups of 1–5 per session, with each participant randomly assigned to treatment group. Each participant was seated at an individual station that included a computer system. First, participants completed the participant questionnaire at their own rates. Second, following instructions that told them to be prepared to answer questions about what was presented, participants received the multimedia program according to their treatment condition. Third, following instructions, participants had 6 min to complete the retention test. Fourth, following instructions, participants had 2.5 min to answer each of the four problem-solving transfer questions, which were presented sequentially. Each test was collected before the next one was distributed. The procedure was the same as the one we used previously (Mayer & Moreno, 1998), except that the multimedia programs involved neutral- and personalized-speech versions.

Scoring. We scored the meteorology knowledge scale by assigning 1 point for each item checked on the 7-item list and assigning an additional 1–5 points corresponding to the participant's self-rating of meteorology knowledge, ranging from 1 point, for *very little,* to 5 points, for *very much.* Participants who scored above 7 were not included in the study or in the count of participants in each group as listed in the *Participants and design* section.

We scored the recall test by tallying how many of eight key-idea units—describing key steps in the process of lightning formation—were included in the participant's recall protocol. The participant did not have to use the same wording as in the multimedia program to receive credit. We scored the problem-solving transfer test by tallying the number of creative solutions generated on the four problem-solving transfer questions, yielding an open-ended scale. Scoring procedures are identical to those we used previously in (Mayer & Moreno, 1998).

Results and Discussion

The primary issue addressed in this study concerns whether students learn the presented material more deeply when the explanations were presented to them in a self-referential personalized conversation. The mean scores and standard deviations of the P and N groups on the transfer and retention tests are presented in Table 1. Students in the P group generated significantly more conceptual creative solutions on the transfer test than did students in the N group, $t(32) = 2.95, p < .01$. Students in the P group did not recall significantly more conceptual idea units on the retention

Table 1
Mean Transfer and Retention Scores, Corresponding Standard Deviations, and Effect Sizes (ES) for the Personalized (P) and Neutral (N) Groups: Experiments 1 and 2

	Group									
	Experiment 1					Experiment 2				
	P		N			P		N		
Score	*M*	*SD*	*M*	*SD*	ES	*M*	*SD*	*M*	*SD*	ES
Transfer	5.87	1.36	4.31	1.62	1.00	5.09	1.69	2.36	1.71	1.60
Retention	5.62	1.02	5.44	1.31	0.15	5.73	1.52	5.41	1.62	0.20

Note. Scores ranged from 0 to 8 for the retention test and from 0 to 10 for the transfer test.

test than did students in the N group, $t(32) = 0.45$, $p = .66$. The effect sizes were 1.00 for transfer and 0.15 for retention. Although both groups remembered equivalent amounts of the verbal explanation for lightning, participants in the P group were better able than those in the N group to use the information to solve novel problems.

Experiment 2

The purpose of Experiment 2 was to determine whether the self-reference effects involving speech in multimedia explanations with animation that occurred in Experiment 1 would also occur with on-screen text messages. In Experiment 2 we studied students' learning from a short multimedia explanation of how lightning storms develop. Animation was accompanied by either neutral or personal on-screen text.

Method

Participants and design. The participants were 44 college students recruited from the psychology subject pool at the University of California, Santa Barbara. There were 22 participants in the personalized-text (P) group and 22 participants in the neutral-text (N) group. As in Experiment 1, all participants reported low levels of experience in meteorology.

Materials and apparatus. The paper materials and apparatus were identical to those used in Experiment 1. The computer programs also were identical to those used in Experiment 1, except the narration was eliminated and converted to on-screen text, which appeared at the bottom of the screen. In this way, the personalized-text program was based on the personalized-speech program used in Experiment 1, and the neutral-text program was based on the neutral-speech program used in Experiment 1.

Procedure. The procedure was identical to that used in Experiment 1 except that the computer programs presented personalized text or neutral text rather than personalized speech or neutral speech.

Scoring. Scoring was identical to Experiment 1. No participant scored above 7 on the meteorology knowledge scale.

Results and Discussion

The mean scores and standard deviations of the P and N groups on the transfer and retention tests are listed in Table 1. Students in the P group generated significantly more conceptual creative solutions on the transfer test than did students in the N group, $t(42) = 5.33$, $p < .001$. Students in the P group did not recall significantly more conceptual idea units on the retention test than did students in the N group, $t(42) = 0.67$, $p = .51$. The effect sizes were 1.60 for transfer and 0.20 for retention. Consistent with Experiment 1, the P group demonstrated better performance on problem-solving transfer than did the N group, but the groups did not differ on tests of retention. The lack of memorial benefits for personalized spoken or written messages seems to be inconsistent with past self-reference effects in memory (Reeder et al., 1987; Symons & Johnson, 1997). A possible interpretation of the difference between our results and results of past self-reference research is that our multimedia explanation provided not only verbal information but also visual information. On the basis of prior multimedia effects (Mayer, 1997), in which students who received text and illustrations or narration and animation performed better on retention tests than students who received text alone or narration alone, we believe that students in both groups used the additional visual information of the animation to help them answer the retention question. However, the major prediction from the interactive hypothesis—that students in the P group would learn more deeply than N students—was confirmed in both studies.

Although the interactive hypothesis can account for the problem-solving transfer results, interest theories of learning (Dewey, 1913; Harp & Mayer, 1998; Renninger, Hidi, & Krapp, 1992) provide an alternative explanation. When students are personally involved in their learning, they are more likely to enjoy the learning situation and actually want to understand the material. When students try hard to make sense of the presented material, they form a coherent mental model that enables them to apply what they learned to challenging new problem-solving situations (Mayer & Wittrock, 1996). For example, personalizing mathematics problems has proven to be more motivational for children (Anand & Ross, 1987; Herndon, 1987; Ross & Anand, 1987). The remaining studies reported in this article are similar to Experiments 1 and 2. To investigate the nature of the resulting learning outcomes, we asked students to take retention and problem-solving transfer tests. In addition, to investigate how interest might affect students' learning, students were asked to rate the program on perceived difficulty, friendliness, helpfulness, motivation, and interest.

Experiment 3

Another example of one-on-one learning scenarios includes a single student playing an educational game on a computer. Just as in the multimedia explanation used in Experiments 1 and 2, a self-reference effect might be equally likely to occur in an educational game that presents instructional messages. However, the level of interactivity and involvement that characterizes an educational game might create in itself the necessary conditions to actively engage students in the learning task. Personalizing messages might not play such a crucial role in this situation. The purpose of Experiment 3 was to examine the self-reference effects within a multimedia game in which students learn about environmental science with the help of a pedagogic agent.

Method

Participants and design. The participants were 39 college students from the psychology subject pool at the University of California, Santa Barbara. Eighteen participants served in the personalized (P) group, and 21 participants served in the nonpersonalized (N) group.

Materials and apparatus. For each participant, the paper-and-pencil materials consisted of a participant questionnaire, a retention test, a 7-page problem-solving transfer test, and a program rating sheet, with each typed on 8.5-in. × 11-in. (21.6 cm × 27.9 cm) sheets of paper.

The participant questionnaire solicited information concerning the participant's name, grade-point average, SAT scores, gender, and knowledge of botany.

The retention test consisted of the following three questions, each typed on the same sheet: (a) "Please write down all the types of roots that you can remember from the lesson," (b) "Please write down all the types of stems that you can remember from the lesson," and (c) "Please write down all the types of leaves that you can remember from the lesson."

The problem-solving transfer test consisted of seven questions, each typed on a separate sheet. The first five sheets, respectively, had the following statements at the top: (a) "Design a plant to live in an environment that has low sunlight," (b) "Design a plant to live in an environment that has low temperature and high water table," (c) "Design a plant to live in an environment that has high temperature," (d) "Design a plant to live in an environment that has heavy rainfall and low nutrients," (e) "Design a plant to live in an environment that has high wind." After each statement, on the same sheet, a diagram with the eight possible roots (branching, deep, thick; branching, deep, thin; branching, shallow, thick; branching, shallow, thin; nonbranching, deep, thick; nonbranching, deep, thin; nonbranching, shallow, thick; nonbranching, shallow, thin), eight possible stems (long, bark, thick; long, bark, thin; long, no bark, thick; long, no bark, thin; short, bark, thick; short, bark, thin; short, no bark, thick; short, no bark, thin), and eight possible leaves (thick, small, thick skinned; thick, small, thin skinned; thick, large, thick skinned; thick, large, thin skinned; thin, small, thick skinned; thin, small, thin skinned; thin, large, thick skinned; thin, large, thin skinned), and their respective names were listed preceded, respectively, by the following text: "Circle the type of root (1 or more):," "Circle the type of stem (1 or more):," and "Circle the type of leaf (1 or more):" Finally, the following question appeared at the bottom of each sheet: "Why do you think that the plant you designed will survive in this environment?"

The following question was on the final two sheets: "In what kind of environment would you expect to see the following plant flourish (i.e., to see the plant grow well)? Please put a check mark next to one or more conditions." Following the question, a diagram of a different plant for each problem was presented in the middle of the sheet, and a list with the eight possible environmental conditions was provided under the diagram (i.e., low temperature, high temperature, low rainfall, heavy rainfall, low nutrients, high nutrients, low water table, and high water table). Problem 6 consisted of a plant with thick, large, and thin-skinned leaves; short, thick, and no-bark stem; branching, shallow, and thin roots. Problem 7 consisted of a plant with thick, small, thick-skinned leaves; thick, long, and bark stem; and nonbranching, deep, and thick roots. Finally, at the bottom of the sheet the following question appeared for each problem: "Why do you think that the plant designed will flourish in the environment that you chose?"

The program ratings sheet contained eight questions asking participants to rate on a 10-point scale their level of motivation, interest, and understanding; the perceived difficulty of the material; and the friendliness of the program. The following question was intended to assess the learner's motivation level: "If you had a chance to use this program with new environmental conditions, how eager would you be to do so?" (rated on a scale that ranged from 1 [*not eager*] to 10 [*very eager*]). The following two questions were intended to assess the learner's interest level: "How interesting is this material?" (scale: 1 [*boring*] to 10 [*interesting*]) and "How entertaining is this material?" (scale: 1 [*tiresome*] to 10 [*entertaining*]). The following two questions were intended to assess the learner's level of understanding: "How much does this material help you understand the relation between plant design and the environment?" (scale: 1 [*not at all*] to 10 [*very much*]) and "How helpful is this material for learning about plant design?" (scale: 1 [*unhelpful*] and 10 [*helpful*]). The following two questions were intended to assess the learner's perception of learning difficulty: "How difficult was the material?" (scale: 1 [*easy*] and 10 [*difficult*]) and "How much effort is required to learn the material?" (scale: 1 [*little*] and 10 [*much*]). The following question was intended to assess the learner's rating of the program's friendliness: "How friendly was the computer that you interacted with?,", and participants had to rate on a 10-point scale the level of friendliness (scale: 1 [*not very friendly*] and 10 [*very friendly*]).

The computerized materials were based on the discovery-based learning environment called "Design-A-Plant" (Lester, Stone, & Stelling, 1999), in which the student travels to an alien planet that has certain environmental conditions (e.g., low rainfall, light sunlight) and must design a plant that would flourish there (e.g., including designing the characteristics of the leaves, stem, and roots). The Design-A-Plant microworld uses a pedagogic agent who offers individualized narrated advice concerning the relation between plant features and environmental features by providing students with feedback on the choices that they make in the process of designing plants. Throughout the program, students visit a total of eight planets. In the first frame of the program, the agent introduces the student to the first set of environmental conditions. Then he asks the student to choose the appropriate root from the library of roots' names and graphics shown on the computer screen. The student chooses a root; if he or she makes a wrong choice, then the agent gives a verbal explanation for the correct root and gives the student a second chance. If the student is wrong again, the agent shows an example of a correct choice and continues with the next step. If the student makes a correct choice, the agent indicates that the choice was correct and proceeds to the next step. The same procedure applies to the stem and leaves, with the pedagogical agent first asking the student to make a choice and giving the student verbal feedback afterward. After the explanation for the appropriate leaves for an environment is given, the student is taken to the next environment. The same procedure follows for the rest of the environments. All explanations are given as speech. Once the last environment is done, the computer screen signals that the program is over.

The two versions of the computerized program for the P group and the N group were identical with the exception of the language style used by the agent. For the P version, a personalized language that characterizes normal conversations was used, with students receiving explanations in the first- and second-person point of view, suggesting that they were sharing the learning experience in the multimedia environment with the agent. For the N version a monologue-style language was used, with students receiving explanations in the third-person point of view, as if they were observers of the multimedia environment that the agent was describing. The factual content of the P and N versions was identical. Appendix B has an example of the speech corresponding to the P and N versions of the lesson. The multimedia programs were developed using Director 4.04 and SoundEdit 16, version 2 (Macromedia, 1995). The apparatus consisted of 3 Macintosh IIci computer systems, which included a 14-in. (35.6-cm) monitor and Sony headphones.

Procedure. Participants were tested in groups of 1–3 per session. Each participant was randomly assigned to a treatment group (P or N) and was seated at an individual cubicle in front of a computer. First, participants completed the participant questionnaire at their own rate. Second, the experimenter presented oral instructions stating that the computer program would teach them how plants should be designed in order to survive in different environments and that when the computer program was finished the experimenter would have some questions for the participants to answer. Students were told to remain quietly seated once the multimedia lesson was over, until the experimenter gave them further instructions. Participants were told to put on headphones and to press the spacebar to begin the program. Third, on pressing the spacebar, the respective version of the multimedia program was presented once to all participants. All participants visited the same eight environments. The self-paced visits had a total duration that ranged between 24 min and 28 min. Fourth, when the program was finished, the experimenter presented oral instructions for the test, stating that there would be a series of question sheets and that for each sheet the participant should keep working until told to stop. Fifth, the program ratings sheet was presented and collected after 2 min. Next, the retention sheet was distributed. After 5 min, the sheet was collected. Then, the seven problem-solving transfer sheets were presented one at a time, for 3 min each, with each sheet collected by the experimenter before the subsequent sheet was handed out.

Scoring. The program rating was calculated by adding the numbers that the participant had circled in the program rating sheet. For this purpose, the perceived difficulty ratings were subtracted from 10. The maximum possible score was 50.

We computed a retention score for each participant by counting the number of correct categories (out of three possible) for each plant part (root, stem, and leaf) that the participant produced on the retention test. The maximum possible score was 9. We computed a total transfer score for each participant by counting the number of acceptable answers that the participant produced across the seven problem-solving transfer problems. For each of the first five questions, 1 point was given for each correct category that the participant circled for each plant part, and 1 point was given for each correctly stated explanation about their choice of category of plant type, regardless of wording. For example, for Question 2, which asked the student to "design a plant to live in an environment that has low temperature and high water table," eight plant categories (branching roots, shallow roots, thick roots, thick stem, bark stem, thick leaves, small leaves, thick skinned leaves) had to be checked, and the student could obtain a maximum possible score of 16: 1 point for each of eight correct categories, plus 1 point for each correct explanation corresponding to the eight possible categories checked. For each of the last two questions, 1 point was given for each correct environment condition chosen by the participant (out of 4), and 1 point was given for each correctly stated explanation about the participant's choice of type of environment, regardless of wording. No points were given for incorrect or incomplete answers. The maximum possible scores for each question were 4 points for Question 1, 16 points for Question 2, 4 points for Question 3, 6 points for Question 4, 12 points for Question 5, and 9 points for each of Questions 6 and 7, yielding a maximum possible total problem-solving transfer score of 60 points.

Results and Discussion

Does personalized speech promote better performance on the transfer and retention tests than neutral speech? The mean scores and respective standard deviations of the P and N groups on retention and transfer are presented in Table 2. Participants in the P group recalled significantly more items than participants in the N group, based on a two-tailed t test, $t(37) = 3.28$, $p < .005$, and produced significantly more correct solutions on problem-solving transfer problems than students in the N group, based on a two-tailed t test, $t(37) = 5.71$, $p < .0001$. The effect sizes were 1.55 for transfer and 0.83 for retention.

Are agent-based programs rated more favorably when they include personalized spoken messages rather than neutral spoken messages? The mean scores and standard deviations of the P and N groups on the overall program rating are presented in Table 2. Groups did not differ on the overall program rating based on a two-tailed t test, $t(37) = 0.69$, $p = .50$. The effect size is 0.21.

The results of Experiment 3 are consistent with the interactive hypothesis. Students who were encouraged to see themselves as interacting with a pedagogical agent in a shared computer-based environment remembered more and used what they learned to solve new problems better than students who were not addressed as participants of the interaction.

Experiment 4

In Experiment 3 we tested the effects of language style in an instructional lesson where students communicated with a pedagogical agent by means of speech. Our goal in Experiment 4 was to investigate whether the same effects can be obtained in an environment where students communicate with the pedagogic agent by reading on-screen text.

Method

Participants and design. The participants were 42 college students from the psychology subject pool at the University of California, Santa Barbara. Twenty-one participants served in the personalized text (P) group, and 21 participants served in the nonpersonalized text (N) group.

Materials and apparatus. The P and N versions of the computer lesson were identical to the P and N versions used for Experiment 3, respectively, with one exception: The voice used in Experiment 3 was replaced by the same words displayed as on-screen text. The text was displayed for the same length of time as the P and N versions of the narration in Experiment 3. Both versions were programmed to have the same duration. The computerized materials were developed using Director 4.04 (Macromedia, 1995). The rest of the materials and the apparatus were identical to those of Experiment 3.

Procedure. The procedure was the same as in Experiment 3 except that each participant was randomly assigned to one of two treatment groups (P or N) before being seated at an individual cubicle in front of a computer.

Scoring. A scorer, not aware of the treatment condition of each participant, determined the retention, problem-solving transfer, and program-rating scores for each participant in the same fashion as in Experiment 3.

Results and Discussion

Does personalized text promote better performance on the transfer and retention tests than neutral text does? The mean scores and respective standard deviations of the P and N groups on retention and transfer are presented in Table 2. Students in the P

Table 2
Mean Transfer, Retention, and Program Ratings Scores, Corresponding Standard Deviations, and Effect Sizes (ES) for the Personalized (P) and Neutral (N) Groups: Experiments 3, 4, and 5

	Experiment and group														
	Experiment 3					Experiment 4					Experiment 5				
	P		N			P		N			P		N		
Score	*M*	*SD*	*M*	*SD*	ES	*M*	*SD*	*M*	*SD*	ES	*M*	*SD*	*M*	*SD*	ES
Transfer	41.89	5.29	28.62	8.54	1.55	46.05	5.43	38.48	4.75	1.58	40.55	5.82	31.86	9.83	0.89
Retention	8.17	0.79	6.67	1.80	0.83	8.19	1.03	7.43	1.36	0.57	8.05	1.17	7.19	1.57	0.56
Program rating	32.67	6.24	31.10	7.76	0.21	32.81	7.38	28.81	5.66	0.70	31.46	5.86	29.41	6.27	0.33

Note. Scores ranged from 0 to 9 for the retention test, from 0 to 60 for the transfer test, and from 5 to 50 for the program rating score.

group recalled significantly more items than students in the N group, based on a two-tailed t test, $t(40) = 2.04$, $p < .05$, and produced significantly more correct solutions on transfer problems than students in the N group, based on a two-tailed t test, $t(37) = 5.71$, $p < .0001$. The effect sizes were 1.58 for transfer and 0.57 for retention.

Are agent-based computer programs rated more favorably when they include personalized textual messages rather than neutral textual messages? The mean scores and respective standard deviations of the P and N groups on the overall program rating are presented in Table 2. Students in the P group gave marginally higher ratings to the computer lesson than students in the N group, based on a two-tailed t test, $t(40) = 1.98$, $p = .06$. The effect size is 0.7.

Similar to Experiment 3, the results obtained in Experiment 4 are consistent with the interactive hypothesis as they confirm its predictions of self-reference effects on retention and transfer. However, at least part of the self-reference effects found in Experiments 3 and 4 might be explained by an alternative interpretation. There is considerable empirical support for the contention that interspersed questions—such as the ones used in the personalized versions of Experiments 3 and 4—result in better learning of both intentional and incidental material (Allington & Weber, 1993; Rickards & McCormick, 1988; Rothkopf, 1970). As can be seen in Appendix B, these personalized messages contained some rhetorical questions interspersed throughout the computer lesson, such as the question "Which do you think would be more flexible?"

If the self-reference effects found in Experiments 3 and 4 were due to the extra elaboration promoted by the inserted questions, then removing the questions from the personalized version would eliminate any performance differences between groups. On the other hand, if the interactive hypothesis is correct, then the benefits of personalized instructional messages rely not on inserted questions but rather on the active involvement resulting from using self-referential language.

Experiment 5

The purpose of Experiment 5 was to study the effects of using self-referential contexts without the inclusion of inserted questions.

Method

Participants and design. The participants were 43 college students from the psychology subject pool at the University of California, Santa Barbara. Twenty-two participants served in the personalized narration (P) group, and 21 participants served in the nonpersonalized narration (N) group.

Materials and apparatus. The P and N versions of the computer lesson were identical to the P and N versions used in Experiment 3, with one exception: The 15 interspersed rhetorical questions used in the P version for Experiment 3 were eliminated. The computerized materials were developed using Director 4.04 and SoundEdit 16, version 2 (Macromedia, 1995). The rest of the materials and the apparatus were identical to those of Experiment 3.

Procedure. The procedure was the same as in Experiment 3.

Scoring. A scorer, not aware of the treatment condition of each participant, determined the retention, transfer, and program-rating scores for each participant in the same fashion as for Experiment 3.

Results and Discussion

Does personalized speech promote better performance on the transfer and retention tests than neutral speech does? The mean scores and respective standard deviations of the P and N groups on retention and transfer are presented in Table 2. Like Experiments 3 and 4, students in the P group recalled more elements from the plant library than students in the N group, based on a two-tailed t test, $t(41) = 2.03$, $p = .05$, and produced significantly more correct solutions on transfer problems than students in the N group, based on a two-tailed t test, $t(41) = 3.55$, $p = .001$. The effect sizes were 0.89 for transfer and 0.56 for retention.

Are agent-based programs rated more favorably when they include personalized spoken messages rather than neutral spoken messages? The mean scores and standard deviations of the P and N groups on the overall program rating are in Table 2. As in Experiments 3 and 4, the groups did not differ on the overall program rating based on a two-tailed t test, $t(41) = 1.11$, $p = .77$. The effect size is 0.33.

Experiment 5 provides evidence that students who communicate with an agent by means of self-referenced messages are able to recall more and use what they have learned to solve new problems better than students who listen to neutral instructional messages. These results are consistent with the interactive hypothesis and suggest that the self-reference effects found in Experiments 3 and 4 were not caused solely by the use of interspersed questions.

In sum, the self-reference effects found across Experiments 3–5 support the use of personalized conversations in student communications with pedagogic agents as a cognitive tool to promote meaningful learning. The transmission hypothesis, according to which students are mere information recipients and the pragmatics of language are not essential, proved to be inadequate by failing to explain the reported effects.

We speculated that performance differences between groups could be explained by interest theories of learning (Dewey, 1913); however, this hypothesis did not receive support in our studies. Nevertheless, it is important to note that our program rating questionnaire may not have been a sensitive enough measure to tap students' interest.

General Discussion

In these experiments we evaluated whether using the self as reference point in a multimedia science lesson can promote deep understanding by engaging students in an active search for meaning (Anderson & Pearson, 1984; Doctorow et al., 1978).

A self-reference effect for retention in multimedia messages was observed in three experiments. On the basis of past self-reference effects on memory (Reeder et al., 1987; Symons & Johnson, 1997), we originally predicted that P groups would recall significantly more of the materials of the computer lesson than N groups in both the multimedia explanation and the multimedia science game. However, this proved to be the case for only the second environment. We interpret the failure to find self-reference effects on retention in the multimedia explanation as due to the additional information contained in the animation. In past research on multimedia learning with animations, students who received scientific explanations by means of text or narration and animations performed better on retention tests than students who received text alone or narration alone (Mayer, 1997). We believe that, congruently, students in Experiments 1 and 2 used the additional visual information of the animation to help them answer the retention question.

A self-reference effect for problem-solving transfer in multimedia messages was observed across five experiments: Students who learned by means of a personalized explanation (either as speech or as on-screen text) were better able to use what they learned to solve new problems than students who received a neutral monologue.

Taken together, the findings fail to support the transmission hypothesis, according to which the pragmatics of language style play no role in the processing of information (Reddy, 1979). The interactive hypothesis, however, was supported across all studies on the dependent measure that is most sensitive to learning: problem-solving transfer. The main focus of our study was to determine if personalized messages would promote the deep understanding of a complex scientific system. Therefore, problem-solving transfer, not retention, should be the most relevant assessment variable (Mayer, 1997; Mayer & Wittrock, 1996).

We hypothesized that students who receive personalized messages would be more likely to attain deep understanding than students who received depersonalized messages. First, addressing students directly in the message and encouraging them to encode the lesson as a personal experience leads to activation of a self-structure. This structure provides internal cues in the form of experiences that facilitate the processing of the content by making available experiences in memory to which the lesson can be related (Burnkrant & Unnava, 1989). Second, by explaining the material in a more familiar and authentic learning style, a one-on-one conversation seems likely to reduce the processing effort required to translate and make sense out of the materials (Mayer, 1984; Pea, 1993). The alternative explanation, based on interest theory (Dewey, 1913), was not supported by our results.

The reported results have strong implications for teaching. The most direct practical implication of the present study is that multimedia science programs can result in broader learning if the communication model is centered around shared environments in which the student is addressed as a participant rather than as an observer. The beneficial effects of introducing self-referencing into a multimedia science lesson occur independently of the *behavioral* interaction required during a computer lesson. When the presentation is linear (such as in Experiments 1 and 2), so that students are required only to watch an animation while listening to or reading an explanation, and when students are required to make choices by clicking on the computer screen (such as in Experiments 3, 4, and 5), self-referencing seems to promote the *mental* interaction needed to actively involve the learner in the process of understanding.

References

Allington, R. L., & Weber, R. (1993). Questioning questions in teaching and learning from texts. In B. K. Britton & A. Woodward (Eds.), *Learning from textbooks: Theory and practice* (pp. 47–68). Hillsdale, NJ: Erlbaum.

Anand, P., & Ross, S. M. (1987). Using computer-assisted instruction to personalize math learning materials for elementary school children. *Journal of Educational Psychology, 79,* 72–79.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. L., Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (Vol. 1, pp. 255–291). White Plains, NY: Longman.

Bretzing, B. H., & Kulhavy, R. W. (1981). Note-taking and passage style. *Journal of Educational Psychology, 73,* 242–250.

Burnkrant, R. E., & Unnava, H. R. (1989). Self-referencing: A strategy for increasing processing of message content. *Personality and Social Psychology Bulletin, 15,* 628–638.

Cherry, E. C. (1953). Some experiments upon the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25,* 975–979.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 11,* 268–294.

d'Ailly, H. H., Simpson, J., & MacKinnon, G. E. (1997). Where should "you" go in a math compare problem? *Journal of Educational Psychology, 89,* 562–567.

Davis-Dorsey, J., Ross, S. M., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problems. *Journal of Educational Psychology, 83,* 61–68.

Dewey, J. (1913). *Interest and effort in education.* Cambridge, MA: Houghton Mifflin.

Doctorow, M., Wittrock, M. C., & Marks, C. (1978). Generative processes in reading comprehension. *Journal of Educational Psychology, 70,* 109–118.

Freedle, R., & Hale, G. (1979). Acquisition of new comprehension schemata for expository prose by transfer of a narrative schema. In R. O. Freedle (Ed.), *New directions in discourse processing* (pp. 121–135). Norwood, NJ: Ablex.

Graesser, A. C. (1981). *Prose comprehension beyond the word.* New York: Springer-Verlag.

Graesser, A. C., Golding, J. M., & Long, D. L. (1998). Narrative representation and comprehension. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 171–205). White Plains, NY: Longman.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90,* 414–434.

Herndon, J. N. (1987). Learner interests, achievement, and continuing motivation in instruction. *Journal of Instructional Development, 10,* 11–12.

Hoban, C. F., & van Ormer, E. (1970). *The literature of the cinema.* New York: Arno.

Lester, J. C., Stone, B. A., & Stelling, G. (1999). Lifelike pedagogical agents for mixed-initiative problem solving in constructivist learning environments. *User Modeling and User-Adapted Interaction, 9,* 1–44.

Macromedia. (1995). Director 4.0.4 [Computer program]. San Francisco: Author.

Mayer, R. E. (1984). Aids to text comprehension. *Educational Psychologist, 19,* 30–42.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32,* 1–19.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90,* 312–320.

Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 47–62). New York: Macmillan.

Mueller, J. H., Bailis, K. L., & Goldstein, A. G. (1979). Depth of processing and anxiety in facial recognition. *British Journal of Psychology, 23,* 511–515.

Pea, R. D. (1993). Learning scientific concepts through material and social activities: Conversational analysis meets conceptual change. *Educational Psychologist, 28,* 265–277.

Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (pp. 284–324). Cambridge, England: Cambridge University Press.

Reeder, G. D., McCormick, C. B., & Esselman, E. D. (1987). Self-referent processing and recall of prose. *Journal of Educational Psychology, 79,* 243–248.

Reeves, B., & Nass, C. (1996). *The media equation.* New York: Cambridge University Press.

Renninger, K. A., Hidi, S., & Krapp, A. (Eds.). (1992). *The role of interest in learning and development.* Hillsdale, NJ: Erlbaum.

Reynolds, C. (1998). As we may communicate. *SIGCHI Bulletin, 30,* 40–44.

Rickards, J. P., & McCormick, C. B. (1988). Effect of interspersed conceptual prequestions on note-taking in listening comprehension. *Journal of Educational Psychology, 80,* 592–594.

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35,* 677–688.

Ross, S. M. (1983). Increasing the meaningfulness of quantitative material by adapting context to student background. *Journal of Educational Psychology, 75,* 519–529.

Ross, S. M., & Anand, P. (1987). A computer-based strategy for personalizing verbal problems in teaching mathematics. *Educational Communication and Technology Journal, 35,* 151–162.

Ross, S. M., McCormick, D., & Krisak, N. (1986). Adapting the thematic context of mathematical problems to student interests: Individual versus group-based strategies. *Journal of Educational Research, 79,* 245–252.

Rothkopf, E. Z. (1970). The concept of mathemagenic activities. *Review of Educational Research, 40,* 325–336.

Spiro, R. (1977). Remembering information from text: The state of schema approach. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 137–165). Hillsdale, NJ: Erlbaum.

Spiro, R. J., & Taylor, B. M. (1987). On investigating children's transition from narrative to expository discourse: The multidimensional nature of psychological text classification. In R. J. Tierney, P. L. Anders, & J. N. Mitchell (Eds.), *Understanding readers' understanding: Theory and practice* (pp. 77–93). Hillsdale, NJ: Erlbaum.

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin, 121,* 371–394.

Turco, R. M. (1996). Self-referencing, quality of argument, and persuasion. *Current Psychology: Developmental, Learning, Personality, Social, 15,* 258–276.

Appendix A

Text of Neutral (N) and Personalized (P) Versions in Experiments 1 and 2, With Personalized Additions Indicated in Brackets and Deletions Indicated by Underlining

[Let me tell you what happens when lightning forms. Suppose you are standing outside, feeling the warm rays of the sun heating up the earth's surface around you.] Cool moist air moves over a warmer surface and becomes heated. The warmed moist air near the earth's surface rises rapidly. As the air in this updraft cools, water vapor condenses into water droplets and forms a cloud. [Congratulations! You have just witnessed the birth of your own cloud!]

[As you watch, you tilt your head skyward. Your] *The* cloud's top extends above the freezing level, so the upper portion of [your] *the* cloud is composed of tiny ice crystals. [Brr! I'm feeling cold just thinking about it!] Eventually, the water droplets and ice crystals become too large to be suspended by updrafts. As raindrops and ice crystals fall through [your] *the* cloud, they drag some of the air in [your] *the* cloud downward, producing downdrafts. When downdrafts strike the ground, they spread out in all directions, producing the gusts of cool wind [you] *people* feel just before the start of the rain. [If you could look inside your cloud, you could see a neat pattern:] *Within the cloud,* the rising and falling air currents cause electrical charges to build. The negatively charged particles fall to the bottom of the cloud, and most of the positively charged particles rise to the top.

[Now that your cloud is charged up, I can tell you the rest of the story:] A stepped leader of negative charges moves downward in a series of steps. It nears the ground. A positively charged leader travels up from objects [around you] such as trees and buildings. The two leaders generally meet about 165 feet above the ground. Negatively charged particles then rush from [your] *the* cloud to the ground along the path created by the leaders. It is not very bright. As the leader stroke nears the ground, it induces an opposite charge, so positively charged particles from the ground rush upward along the same path. This upward motion of the current is the return stroke. It produces the bright light that [you] *people* notice as a flash of lightning.

Appendix B

Text of Neutral (N) and Personalized (P) Versions in Experiments 3 and 4 Corresponding to the Introduction and First Environment Visit

Introduction to the Program

P version: You are about to start a journey where you will be visiting different planets. For each planet, you will need to design a plant. Your mission is to learn what type of roots, stem, and leaves will allow your plant to survive in each environment. I will be guiding you through by giving out some hints.

N version: This program is about what type of plants survive in different planets. For each planet, a plant will be designed. The goal is to learn what type of roots, stem, and leaves allow plants to survive in each environment. Some hints are provided throughout the program.

Introduction to the Environment

P version: Your only goal here is to design a plant that will survive, maybe even flourish, in this environment of heavy rain. It is perfect for any of the roots and stems, but your leaves need to be flexible so that they won't be damaged by the heavy rain.

N version: The goal is to design a plant that will survive, maybe even flourish, in an environment of heavy rain. It is perfect for any root and stem, but the leaves need to be flexible so that they won't be damaged by the heavy rain.

Leaf Explanation

P version: This is a very rainy environment, and the leaves of your plant have to be flexible so they're not damaged by the rainfall. What really matters for the rain is your choice between thick leaves and thin leaves. Which do you think would be more flexible?

N version: In very rainy environments plant leaves have to be flexible so that they are not damaged by the rainfall. What really matters for the rain is the choice between thick leaves and thin leaves.

Feedback

Right Root Choice

P version: Yes! In this environment any root you choose will do just fine.

N version: For a heavy rain environment any root will do just fine.

Right Stem Choice

P version: Yes! In this environment any stem you choose will do just fine.

N version: For a heavy rain environment any stem will do just fine.

Right Leaf Choice

P version: Yes! You chose thin leaves, which are very flexible and won't break in the rain.

N version: The correct choice is thin leaves, which are very flexible and don't break in the rain.

Wrong Leaf Choice

P version: Choose thin leaves!

N version: The right choice is thin leaves.

Received December 29, 1999
Revision received April 29, 2000
Accepted April 29, 2000 ■