Pedagogical Agents in Constructivist Multimedia Environments: The Role of Image and Language in the Instructional Communication

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

Roxana Moreno
Department of Psychology
University of California
Santa Barbara, CA 93106

Objectives and Theoretical Framework

What are the cognitive consequences of introducing the image of a pedagogical agent in a discovery-based multimedia learning environment? How should pedagogical agents communicate with students to promote constructivist learning? In order to answer these questions, we conducted two sets of studies. First, we varied whether the agent's words were presented as speech or on-screen text and whether or not the agent's image appeared on the screen, both with an animated fictional agent (Experiment 1), and a video of a human face (Experiment 2). Next, we varied whether or not the agent's words were presented in a self-referenced or neutral style (i. e. as dialogue or monologue) both using speech (Experiment 3), and on-screen text (Experiment 4).

The dissemination of software agents in instructional design brings up the need to investigate the effects of agents' presence in human-computer interaction. Software pedagogical agents may carry the modeling, scaffolding, and coaching attributes of successful mentors in apprenticeship learning (Collins, Brown, & Newman, 1989). However, in addition to the considerations about what type of advise and feedback the agent should give to the student, it is necessary to investigate how the agent should look and talk to the student to convey effectively his message. For example, as mentors are usually physically present during the apprenticeship interaction, software agents may carry the attribute of visual presence (Towns et al., 1998). In addition, as mentors do not merely send out messages into the void to be picked up by students but rather interact with students in the process of knowledge construction, software agents may carry the attribute of human communication (Tharp & Gallimore, 1991).

Regarding the image and voice of the personal agent, we tested the predictions arising from two opposing hypothesis: the social-cue hypothesis, and the interference hypothesis. According to the former, students learn better and rate more favorably computer lessons that include social cues--such as facial expressions or human voices than those which do not include social cues. On the other hand, according to the interference hypothesis, social cues may be entertaining but interfere with the student's ability to make sense out of the presented materials (Moreno & Mayer, in revision-a; Harp & Mayer 1998). Given the constraints on the learner's working memories, instructional programs which include the additional information contained in the agent's image or voice--called seductive details--hinder learning (Moreno & Mayer, in press).

Regarding the language style used by the pedagogical agent, we tested the opposing predictions from the transmission hypothesis and the interactive hypothesis. According to the former, human communication involves encoding an idea into a signal by a sender, transmitting the signal to the receiver, and decoding the signal by the receiver (Reddy, 1979). Consequently, as the content material is identical, changing the agent's language style from a personalized dialogue to a neutral monologue does not affect students' learning. In contrast, the interactive hypothesis claims 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 (Moreno & Mayer, in revision-b). This hypothesis predicts that students will learn deeper from self-referenced messages rather than from neutral messages.

Methods and Data Source

We conducted a total of four studies where students learned the relation between the physical characteristics of a plant and its ability to function in various environments through a multimedia lesson. All students were recruited from the Psychology Subject Pool at the University of California, Santa Barbara. In the first study, 17 students learned by interacting with the image of a fictional agent who spoke to them (Group IN), 16 students learned by interacting without the image of a fictional agent who spoke to them (Group -IN); 15 students learned by interacting with the image of a fictional agent who gave explanations as on-screen text (Group IT); and 16 students learned by interacting without the image of a fictional agent who gave explanations as on-screen text (Group -IT). The second study included the same treatment conditions but the fictional agent was replaced by the video and voice of a human agent. 19 students participated in the IN group, 20 students participated in the -IN group, 19 students participated in the IT group, and 21 students participated in the -IT group.

In the third study, 18 students learned with personalized messages spoken by the agent (Group P), and 21 students learned with a neutral messages spoken by the agent (Group NP). For the fourth study, 21 students learned with a personalized messages displayed as text (Group P), and 21 students learned with neutral messages displayed as text (Group NP).

For each study, the participants were randomly assigned to a treatment group and seated at an individual cubicle in front of a computer. First, students were given a questionnaire, which solicited participant's demographic information. Second, the respective version of the multimedia program was presented. All participants visited eight different environments and had to design the right plant to flourish under the presented conditions. The multimedia program used is called "Design-a-Plant", and was developed by the Multimedia Laboratory at the Department of Computer Science of North Carolina State University (Lester et al 1997). It includes a lifelike pedagogical agent who provides advice to learners as they graphically assemble plants from a library of plant structures such as roots, stems and leaves.

After the respective program was over, all subjects were given a retention test, a 7-page problem-solving test, and a program-ratings sheet. In the retention test, students had to name as many types of roots, stems, and leaves as they could. In the problem-solving test, students had to either check one of the possible kinds of roots, stems and leaves that corresponded to a given environment (first five questions), or check all the possible environmental conditions that caused a given plant to grow well (last two questions). Students were also asked to write an explanation of the choices. The program-ratings sheet asked participants to rate on a 10-point scale their level of motivation, interest, understanding, and the perceived difficulty of the material. We determined for each study, whether the groups differed on measures of retention, transfer, and program-ratings.

Results

Do students who communicate with a pedagogical agent via speech show deeper understanding from a multimedia science lesson than students who communicate with a pedagogical agent via on-screen text? The findings from Experiments 1 and 2 gave evidence in favor of students' communicating with a fictional agent by means of speech by demonstrating what we have called a modality effect in program ratings, recall, and transfer: Students who learn with the voice of an agent rate the lesson more favorably, recall more, and are better able to use what they have learned to solve problems than students who learn the same verbal materials as on-screen text. For Experiment 1, the mean program ratings for narration and text groups was $\underline{M}$s = 24.94 and 22.36, and $\underline{SD}$s = 5.06 and 4.86, respectively, $\underline{F}(1, 60) = 4.19$, $\underline{MSE} = 103.11$, $\underline{p} < 0.05$; the mean number of ideas recalled for narration and text groups was $\underline{M}$s = 8.12 and 7.10, and $\underline{SD}$s = 0.96 and 1.70, respectively, $\underline{F}(1, 60) = 9.30$, $\underline{MSE} = 17.30$, $\underline{p} < .005$; and the mean number of correct answers for narration and text groups was $\underline{M}$s =39.09 and 31.20, and $\underline{SD}$s = 6.82 and 8.85, respectively, $\underline{F}(1, 60) = 16.16$, $\underline{MSE} = 1000.78$, $\underline{p} < .0005$. For Experiment 2, the mean program ratings for narration and text groups was $\underline{M}$s = 31.34 and 29.51, and $\underline{SD}$s = 6.63 and 7.85, respectively, $\underline{F}(1, 75) = 3.93$, $\underline{MSE} = 203.07$, $\underline{p} = 0.05$; the mean number of ideas recalled for narration and text groups was $\underline{M}$s = 8.10 and 7.30, and $\underline{SD}$s = 0.82 and 1.49, respectively, $\underline{F}(1, 75) = 8.58$, $\underline{MSE} = 12.76$, $\underline{p} < 0.005$; and the mean number of correct answers for narration and text groups was $\underline{M}$s =39.95 and 28.40, and $\underline{SD}$s = 6.35 and 8.41, respectively, $\underline{F}(1, 75) = 46.70$, $\underline{MSE} = 2604.44$, $\underline{p} = .0001$.

<u>Do students who are presented with the image of a pedagogical agent show deeper understanding from a multimedia science lesson than students who are not presented with the image ?</u> The findings from Experiments 1 and 2 did not provide evidence in favor or against presenting students with the image of a pedagogical agent, failing to confirm what we have called an <u>image effect</u> in program ratings, recall, and transfer: Students who are presented with the image of an agent do not rate the lesson more favorably, recall more, or are better able to use what they have learned to solve problems than students who are not presented with the visual presence of the agent.

<u>Do students who communicate with a pedagogical agent via a personalized dialogue show deeper understanding from a multimedia science lesson than students who communicate with a pedagogical agent via an non-personalized monologue?</u> The findings from Experiments 3 and 4 gave evidence in favor of students' communicating with a pedagogical agent by means of a personalized conversation by demonstrating what w have called a <u>self-reference effect</u> for recall and retention: Students who learn by communicating with a pedagogic agent via a self-referenced message recall more and are better able to use what they have learned to solve problems than students who learn via a neutral message. For Experiment 3, the mean number of ideas recalled for personalized and neutral groups was $\underline{M}$s = 8.17 and 6.67, and $\underline{SD}$s = 0.79 and 0.39, respectively, $\underline{t}(37) = 3.28$, $\underline{p} < .005$; and the mean number of correct answers for personalized and neutral groups was $\underline{M}$s =41.89 and 28.62, and $\underline{SD}$s = 5.29 and 8.54, respectively, $\underline{t}(37) = 5.71$, $\underline{p} < .0001$. For Experiment 2, the mean number of ideas recalled for personalized and neutral groups was $\underline{M}$s = 8.54 and 7.43, and $\underline{SD}$s = 1.03 and 1.36, respectively, $\underline{t}(40) = 2.04$, $\underline{p} < .05$.; and the mean number of correct answers for personalized and neutral groups was $\underline{M}$s = 46.05 and 38.48, and $\underline{SD}$s = 5.43 and 4.75, respectively, $\underline{t}(37) = 5.71$, $\underline{p} < .0001$. However, the two experiments failed to reveal a self-reference effect for interest.

<u>Scientific and Educational Significance</u>

The reported results have important theoretical and practical implications. First, the modality effect found in Experiments 1 and 2 might be explained as a combination of the superiority in recall when words are processed auditorily rather than visually (Moreno &

Mayer, 1999) and the extra emotional cues which voices carry and text lacks. Congruent to the social-cue hypothesis, voices are powerful indicators of social presence and its incorporation in the interaction might promote richer processing by the incorporation of the additional attitudes and beliefs that are attached to the agent (Reeves & Nass, 1996). Second, the failure to find an image effect in Experiments 1 and 2 does not support either the social-cue hypothesis or the interference hypothesis. However, the conclusions drawn from our results are limited by the materials used in the computer-based lesson. Due to the scientific nature of the lesson, the facial expressions may not have offered any informational advantage to the students. Third, the self-reference effects found in Experiments 3 and 4 support the interactive hypothesis and are consistent with prior self-reference effects on memory and comprehension (Symons & Johnson, 1997). Providing explanations by means of self-referenced dialogues rather than depersonalized generic monologues encourages the learner to actively search for meaning (Moreno & Mayer, in review-b).

The most direct practical implication of these studies is that in a constructivist science lesson where students learn with the help of a pedagogical agent, the agents' voice and conversational style play a fundamental role in the promotion of meaningful learning.

References

Collins, A., Brown, J. S., & Newman, S. E.(1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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.

Lester, J., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. (1997). The persona effect: Affective impact of animated pedagogical agents. In *Proceedings of CHI'97* (pp. 359-366).

Moreno, R. & Mayer, R. E. (in review-a). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction.*

Moreno, R. & Mayer, R. E. (in review-b). Engaging Students in Active Learning: The Case for Personalized Multimedia Messages. *Journal of Educational Psychology.*

Moreno, R. & Mayer, R. E. (in press). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology.*

Moreno, R. & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity effects. *Journal of Educational Psychology, 91,* 1-11.

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

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

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

Tharp, R. G., & Gallimore, R. (1991). *The instructional conversation: Teaching and learning in social activity.* Report from the Office of Educational Research and Improvement. U. S. Department of Education.

Towns, S. G., Callaway, C. B., Voerman, J. L., and Lester, J. C. (1998). Coherent gestures, locomotion, and speech in life-like pedagogical agents. In *Proceedings of the Fourth International Conference on Intelligent User Interfaces*. San Francisco.

Note

This research was supported by a grant from the National Science Foundation entitled "Learning and Intelligent Systems: Animated Pedagogical Agents for Constructivist Learning Environments."