Life-Like Pedagogical Agents in Constructivist Multimedia Environments: Cognitive Consequences of their Interaction

Roxana Moreno
Psychology Department
University of California
Santa Barbara, CA 93106
moreno@ psych.ucsb.edu

Richard E. Mayer
Psychology Department
University of California
Santa Barbara, CA 93106
mayer@ psych.ucsb.edu

James C. Lester
Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206
lester@eos.ncsu.edu

Abstract: The goal of this study is to test the hypothesis that animated pedagogical agents can promote constructivist learning in a discovery-based learning environment. We do this by first, comparing the learning outcomes of students who learn in the context of social interaction with a computer-mediated character with the learning outcomes of students who learn in a computer text context. Second, we focus on the relative contributions of pedagogical agents' image, voice, and language style in promoting deeper learning.

The Case for Software Pedagogical Agents

Discovery-based learning environments are intended to facilitate constructivist learning through creative problem-solving experiences. However, the characteristics of discovery-based environments which offer the greatest potential for fostering learning also pose the greatest challenge: the complexity of the learning experience, based on the learner's freedom to explore and design artifacts in microworlds, constantly threatens to overwhelm learners. To combat this complexity, discovery-based learning environments should provide scaffolding in the form of highly contextualized problem-solving advice that is customized to each learner. Perhaps the most intriguing vehicle for providing such dynamically individualized scaffolding is the emerging technology of animated pedagogical agents--lifelike on-screen characters who provide contextualized advice and feedback throughout a learning episode (Lester & Stone 1997, Lester et al. 1997).

A Life-Like Pedagogical Agent Example

The discovery-based learning environment used in this study is a microworld called "Design-A-Plant", developed by the Multimedia Laboratory at the Department of Computer Science of North Carolina State University (Lester & Stone 1997, Lester et al. 1997). In this program, 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). It includes an animated pedagogic agent called Herman--a lovable alien with appealing facial expressions, human-like movements, and an amusing voice, who offers individualized advice concerning the relation between plant features and environmental features, encouragement when students encounter difficulties, and feedback on the choices that students make in the process of designing plants. Herman's actions are dynamically selected and assembled by a behavior sequencing engine that guides the presentation of problem-solving advice to learners, as described more fully in Lester et al. (1997). The program starts with Herman introducing 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. After the student had chosen a root, Herman gives two explanations: first, a narrated explanation, and second a QuickTime movie that contains a similar explanation for the correct root. The same procedure applies to the stem and leaves, with Herman first asking the student to make a choice, and giving the student feedback afterwards. Once the student was given the last explanation on the leaves for each environment, he is taken to the next environment. The same procedure follows for the rest of the environments.

Issue 1: Agent versus Text-Based Environments

Does learning with an animated pedagogical agent in a discovery environment promote constructivist learning? In order to help answer this question, we conducted a preliminary study where the learning outcomes of students who learned about environmental science in the Design-A-Plant microworld (personal agent or PA group) was compared with the learning outcomes of students who received the identical verbal and visual materials in a computer-based text environment (no personal agent or no-PA group).

Method and Results

The participants were 44 college students from the Psychology Subject Pool at the University of California, Santa Barbara. Twenty-four participants served in the no-PA group and twenty participants served in the PA group. The computerized materials consisted of two multimedia computer programs on how to design a plant. The PA version was the original multimedia program called "Design-a-Plant", where students see a library of plant parts, pick the plant part that they consider appropriate for the respective environment, and receive spoken feedback in a conversational style from the agent (Lester & Stone 1997, Lester et al. 1997). The no-PA version was presented with the same library plant parts and explanations than the PA version, but Herman's image was deleted. In addition, students in the no-PA version were not allowed to design the plant before reading the verbal explanations, but rather received the explanation directly in a monologue style, similar to when science material is read from a text book.

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. After the respective program was over, participants were tested on three important measures of learning: retention--in which we assessed memory for the basic factual information that was presented; problem-solving transfer--in which we asked students to solve new problems based on the principles learned in the respective program; and motivation and interest--in which we asked students to rate their level of motivation, interest, understanding, and the perceived difficulty and friendliness of the lesson. We determined whether the groups differed on measures of retention, transfer, and self-ratings.

Do students who learned interactively with a pedagogical agent show deeper understanding from a multimedia science lesson than students who learned in a conventional environment?

Students in the PA and no-PA groups received identical factual information consisting of the same verbal and visual materials. As a retention test, we simply asked students to name as many types of roots, stems, and leaves as they could--highly relevant information that was prominently displayed in both instructional programs. The mean number of correctly recalled items by the PA group was not significantly different than the mean number of correctly recalled items by the no-PA group. The results suggest that when retention of factual information is the goal of instruction, then environments, which allow for interacting with an animated pedagogical agent are not warranted. Although these results demonstrate that both groups learned the basic factual information, the major prediction in this study is that students in the PA group will learn more deeply than no-PA students. As predicted, students in the PA group produced significantly more correct solutions on transfer problems than students in the no-PA group ($p < .005$).

The foregoing results are consistent with the idea that students who learn with animated pedagogic agents work harder to make sense of the material than do students who learn in a more conventional text-based environment. Another way to test this idea is to assess students' motivation to continue learning and their interest in the material, so we predicted that the PA group would produce a higher mean program rating than the no-PA group. As predicted, the PA group rated their motivation to continue learning and their interest in the material significantly higher than the no-PA group ($p < .01$).

Conclusion

This preliminary study contributes to multimedia learning theory (Mayer 1997) by comparing the cognitive consequences of students' experience with or without the aid of a pedagogical agent. Students in both groups received identical verbal and visual material presented in the context of social interaction with a computer-mediated character (PA group) or in an non-personalized text context (no-PA group). These findings give preliminary evidence in favor of using pedagogical agents as software mentors, and demonstrate a *personal-agent effect* in multimedia learning environments: Students are more motivated, interested, and achieve better transfer when the lesson is imparted by a pedagogic agent rather than by on-screen non-personalized text.

How do agents affect student learning? Reeves and Nass (1996) have provided convincing evidence that students view their contact with computer-based characters as social interactions. Congruent with this approach, students' learning with the pedagogical agent could have been promoted by at least three social-cues: the pedagogical agents' image, voice, and personalized language style. Therefore, we investigated first, the role of pedagogical agents' auditory and visual presence in a discovery-based computer lesson. For this purpose, 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). Second, we investigated the role of the agents' language style in the computer lesson. For this purpose we varied whether or not the agent's words were presented in conversational style (i. e. as dialogue or monologue) both using speech (Experiment 3), and on-screen text (Experiment 4).

Issue 2: The Role of Life-Like Pedagogical Agents' Visual and Auditory Presence

We tested the predictions arising from two opposing hypothesis: the social-cue hypothesis, and the cognitive load hypothesis. According to the former, students communicate better, become more interested, and therefore 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 (Dewey 1913, Rutter 1984). Three predictions can be made for students who learn a multimedia lesson with the image of an agent. First, they will be more likely to remember the materials of the lesson than students who are not provided with the visual presence of the agent; this we call an image effect on retention. Second, they will be more likely to use what they have learned to solve problems than students who are not provided with the visual presence of the agent; this we call an image effect on transfer. Third, they will rate the lesson more favorably as expressed in the program ratings than students who are not provided with the visual presence of the agent; this we call an image effect on program rating. In contrast, according to a cognitive load of learning (Chandler & Sweller 1991, Moreno & Mayer in press, Sweller 1989), the inclusion of irrelevant adjuncts in a science lesson will divert the limited cognitive resources available for the processing of the relevant materials of the lesson. As a result, learning and problem solving will be impaired.

An additional goal of the present study was to investigate the role of pedagogical agents' auditory presence in a constructivist multimedia lesson. We investigated if prior findings on modality effects in learning--where students learn better from visual and verbal presentations when the verbal information is presented as speech rather than as on-screen text, extend to discovery-based multimedia environments (Mayer & Moreno 1998, Moreno & Mayer 1999, Mousavi, Low, & Sweller 1995). Congruent with past research on modality effects in multimedia learning, two predictions can be made. First, students who learn with the voice of an agent will be more likely to remember the materials of the lesson than students who learn the same verbal materials as on-screen text; this we call a modality effect on retention. Second, students who learn with the voice of an agent will be more likely to use what they have learned to solve problems than students who learn the same verbal materials as on-screen text; this we call a modality effect on transfer. An additional prediction was made based on the social-cue hypothesis (Dewey 1913). Students who learn with the voice of an agent will be more motivated and will rate the lesson more favorably than students who learn the same verbal materials as on-screen text; this we call a modality effect on program rating.

Method and Results

In the first study, 17 students learned by interacting with the image of a life-like fictional agent who spoke to them (Group IN), 16 students learned by interacting without the image of an agent who spoke to them (Group -IN); 15 students learned by interacting with the image of a life-like fictional agent who gave explanations as on-screen text (Group IT); and 16 students learned by interacting without the image of an agent who gave explanations as on-screen text (Group -IT). The second study was identical but the video and voice of a human agent replaced the fictional agent. Nineteen 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. The procedure was identical to that of the preliminary study.

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 ?

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 *image effect* in program ratings, recall, and transfer: Students who are presented with the image of an agent do not rate the lesson more favorable, 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.

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 pedagogical 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. In both experiments, the mean program ratings for the narration groups was significantly higher than the mean program ratings for the text groups ($p < 0.05$ for both experiments); the narration groups recalled significantly more than the text groups ($p < .005$ and $p < 0.005$ for Experiments 1 and 2, respectively); and the narration groups gave significantly more correct answers in the transfer tests than the text groups ($p < .0005$ and $p = .0001$ for Experiments 1 and 2, respectively). These results extend the modality effect in learning from visual and verbal materials to discovery-based multimedia environments and confirm the social-cue prediction according to which students are more motivated and interested in programs, which communicate via speech.

Issue 3: The Role of Life-Like Pedagogical Agents' Language Style

Regarding the language style used by the pedagogical agent, we tested the opposing predictions from the transmission hypothesis and the conversational 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 when the language style is changed from a dialogue style to a monologue style, it should not affect students' learning. In contrast, the conversational hypothesis views communication as inherently cooperative and claims that a uni-directional model of communication such as the one supported by the transmission hypothesis, fails to capture the social and individual processes involved in knowledge construction through conversations (Brennan 1990).. This hypothesis predicts that students will learn deeper from instructional conversations rather than from non-personalized monologue-style communications.

Method and Results

In the third study, 18 students learned with a personalized conversation spoken by the human agent used in Experiments 1 and 2, (Group P) and 21 students learned with a non-personalized monologue spoken by the same human agent (Group NP). For the fourth study, 21 students learned with a personalized conversation displayed as text (Group P) and 21 students learned with a non-personalized monologue displayed as text (Group NP). In both studies the image of the agent was not presented. The procedure was identical to that of the prior studies.

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?

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 *dialogue effect* in recall and transfer: Students who learn by communicating with a pedagogic agent via a personalized dialogue recall more and are better able to use what they have learned to solve problems than students who communicate via a non-personalized monologue. The mean number of ideas recalled for dialogue groups was significantly larger than for monologue groups ($p < .005$ and $p < .05$ for Experiments 3 and 4, respectively) and the mean number of correct answers in the transfer test was significantly larger for the dialogue groups than for the monologue groups ($p < .0001$ for both experiments). In addition, Experiment 4 demonstrated a dialogue effect for program ratings. Students who learned by reading dialogue-style text rated more favorably the lesson than students who learned by reading monologue-style text ($p = 0.05$).

General Discussion

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 (Penney 1989), the relatively effortless maintenance of the auditory input in comparison to the visual input provided by text (Anderson & Craik 1974), and the extra emotional cues, which voices carry and text lacks. According to communications research, 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). These findings extend pioneering demonstrations of modality effects in learning from visual and verbal materials (Mayer & Moreno 1998, Moreno & Mayer 1999, Mousavi, Low, & Sweller 1995) in three ways: (a) by examining a modality effect in an interactive constructivist computer-based multimedia environment rather than a paper-based or non-interactive computer environment, (b) by employing multiple dependent measures including students' rating of the learning materials, and (c) by using fictional and non-fictional pedagogic agents to deliver the verbal materials of the lesson. The finding that students rated the computer program more favorably when it communicates via speech rather than via text is congruent with the social-cue hypothesis according to which the agents' voice may help students feel an emotional connection with the agent therefore promoting interest in the learning task which in turn fosters constructivist learning .

Second, the failure to find an image effect in Experiments 1 and 2 does not support the social cue hypothesis or the cognitive load hypothesis. Several interpretations for this finding can be offered. First, the lack of image effects might have relied on the strength of the auditory cues given by the voice of the agents which in both applications was extremely expressive and clear. When a voice carries these qualities, it is less likely that facial expressions or lip movements will provide extra information, or help disambiguate the message in the lesson. An alternative explanation may reside on the specific content of the computer lesson: The scientific nature of the lesson used in the reported studies is not essentially emotional and facial expressions may not offer any informational advantage.

Cognitive load predicted that the introduction of the agent's image in the computer lesson would be detrimental to students' learning. This prediction was not confirmed. However, in our studies, the agent's image was never presented simultaneously with other visual materials of the lesson. According to cognitive load theory, a detrimental effect in learning occurs in the cases that students need to split their attention between mutually referring materials. Therefore, it is more likely that if the multimedia lesson contained animations of the agent presented simultaneously with graphics or text, his visual presence would have been detrimental rather than neutral to learning (Mayer & Moreno 1998, Moreno & Mayer 1999, Sweller 1989).

The dialogue effects on retention and transfer found in Experiments 3 and 4 support the conversational hypothesis. We attribute these effects to a combination of less cognitive effort and more active processing. First, it might be argued that students who listen to non-personalized explanations need to work harder than students who listen to personalized explanations because they need to map the agent's generalized explanations to their specific choices in the computer lesson. Second, conversations might act as attention enhancers and produce more active processing of the materials (Reeves & Nass 1996) by creating a state of vigilance in which the participant might be asked to intervene in the dialogue at any time.

It is tempting to conclude that the personal agent effect found in our preliminary study resided in the effectiveness of teaching via instructional conversations (Tharp & Gallimore 1991). However, our focus in the present study was solely on the role that the agent's social cues may have in a multimedia constructivist environment. Although speech and language style proved to help students understanding, other factors, such as students' interaction with the computer program, might play an important role in learning as well. More research is needed to determine the contribution of alternative interactive designs in promoting deep learning from constructivist environments.

The most direct practical implication of our studies is that in a constructivist science lesson where students learn with the help of a pedagogical agent, technological advances in education should focus on extending the capabilities of speech recognition and natural language systems to facilitate instructional conversations. Moreover, the reported findings suggest that an agent interface metaphor might already be present once students interact with a computer, and trying to make it more visually apparent, may not necessarily lead to better results.

References

Anderson, C. M. B., & Craik, F. I. M. (1974). The effect of a concurrent task on recall from primary memory. Journal of Verbal Learning & Verbal Behavior, 13, 107-113.

Brennan, S. E. (1990). Conversation as direct manipulation. In Laurel, B. (Ed.)., *The art of human-computer interface design.* Reading, MA: Addison-Wesley.

Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction. Cognition and Instruction, 8, 293-332.

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

Lester, J. C. and Stone, B. A., (1997). Increasing believability in animated pedagogical agents. In Proceedings of The First International Conference on Autonomous Agents (pp. 16-21). New York, NY: ACM Press.

Lester, J. C., 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). New York, NY: ACM Press.

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.

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.

Mousavi, S., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*, 319-334.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition, 17*, 398-422.

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.

Rutter, D. R. (1984). Looking and Seeing: the role of visual communication in social interaction. John Wiley & Sons, Chichester.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. Journal of Educational Psychology, 81, 457-466.

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.

Acknowledgments

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