Influence of Colearner Agent Behavior on Learner Performance and Attitudes

Wendy Ju, Seth Nickell, Katherine Eng, Clifford Nass
Dept. of Communication, Stanford University
Stanford, CA 94305-2050
wendyju@stanford.edu, snickell@stanfordalumni.org, keng@stanfordalumni.org, nass@stanford.edu

ABSTRACT
This study examines the effect of colearner agent performance and social behavior on learner performance and subjective satisfaction in an interactive learning environment. In this 2 (high- or low-performing colearner) by 2 (socially supportive or competitive colearner) experiment (N=44), participants learned Morse Code alongside an agent colearner. Participants with high-scoring colearner agents performed significantly better than participants with low-scoring colearners. Participants liked and felt liked by socially supportive agents more than they did socially competitive agent participants. Implications for developing educational software are discussed.

Author Keywords
Colearners, Interactive Learning, Intelligent Tutoring Systems, Agents, Multimedia Information Systems

ACM Classification Keywords
H.5.1. Multimedia Information Systems: Animations, Evaluation/methodology

INTRODUCTION

Interactive Learning Environments create educational settings for users anywhere that they might have access to a computer, a personal digital assistant, or even a cell phone. Inexpensive and portable platforms for educational software are being used not only to supplement grade school education, but also to provide distance learning, to enable on-the-job training, and to allow for constructive entertainment.

One notable shortcoming of these interactive learning environments is that users give up the immediacy of the classroom environment for the opportunity to learn anytime and anywhere. While computers are hard-pressed to replicate the classroom experience, simulation of key aspects of traditional classroom environments can help to improve knowledge acquisition.

Typical classroom learning environments include multiple students who exhibit a wide array of personalities and behaviors. This heterogeneous group of *colearners* influences how individual students learn and how they feel about the learning process. However, research into the effects of colearners on the individual have been limited because it would be virtually impossible to select and manage the classroom environment to enhance each individual student's learning experience.

Research in interactive agents by Reeves and Nass [9] and by Lester [5] suggests that we may respond to the behaviors of animated colearner agents in a software learning environment in much the same way that we respond to human colearners in a traditional classroom. Unlike human colearners, however, colearner agents can be programmed to enhance the student learning experience. Thus, it is possible and highly desirable to ask: *What is the effect of the behavior and performance of colearner on the student learning experience?*

PRIOR WORK

Educational Performance

Research in the field of educational psychology has long found that cooperative study is superior to individual study [6]: children collaborating with other children of differential skill levels in a Lego building task learn more than students working alone [1]; students who learn a topic in social conditions tend to perform better in post-test measures than students learning in individualistic conditions [4]; students acquire factual material more easily when studying with a partner rather than alone [12]. In these studies, the ability to determine which aspects of cooperative or collaborative learning were driving the performance improvements (e.g., mere presence, shared knowledge, attitudinal changes) was limited by the inherent heterogeneity of human colearners.

CHI 2005, April 2–7, 2005, Portland, Oregon, USA.
ACM 1-59593-002-7/05/0004.

Colearner Agents

The predominant theories of collaborative learning—socio-constructivist [8], socio-cultural [11], socially-distributed cognition [3]—presume that the improvement is the result of some shared reasoning or dialogue that takes place between the peer learners. Common experience indicates, however, that our fellow students have a major influence on our learning experience even when their cognitive processes are not transparent to us at all. We are easily affected by knowing that someone is getting all the answers right or wrong, or that the student next to us is rooting for us or against us. That is, *colearners may influence learning not only through their knowledge but also through their behaviors*.

Because it is challenging to create a controlled set of colearners in a real classroom environment, no experimental data exists to prove or disprove such a hypothesis. Hence, the advent of colearner agents provides a key opportunity. However, prior work in the realm of colearning agents, such as that of Chan and Baskin [2], focuses on affecting student performance through advances in instrinsic agent intelligence rather than on demonstrated agent performance.

METHOD

We performed a two by two laboratory experiment during which participants interacted with a Morse Code tutorial program that had an avatar which represented the user, a teacher agent, and a colearner agent. Half of the participants had a colearner that made socially competitive comments (n= 22), whereas the other half (n= 22) had a colearner that made socially supportive statements. In each of the previous conditions, half of the colearners showed few errors during the in-class quizzes, while half made many errors.

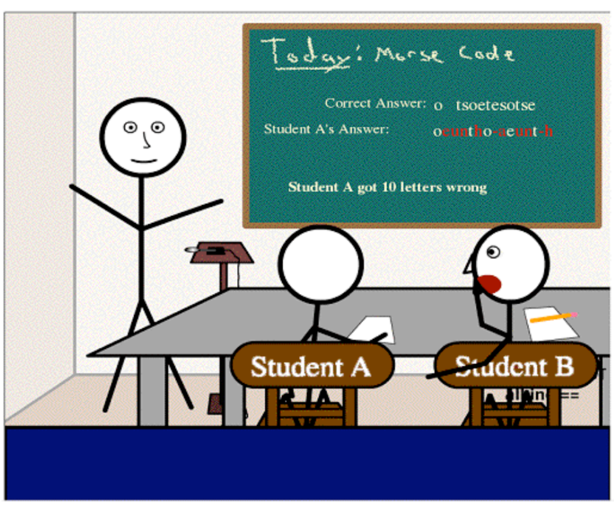

Figure 1. The Virtual Classroom

Participants

44 adults participated in the experiment. Participants were screened prior to the experiment for knowledge of the material being taught in the experiment. Randomization and gender balancing was achieved by assigning participants to conditions round-robin based on their gender and order of arrival to the testing site. All computers used in the experiment were of the same make and model.

Procedure

Each participant was assigned their own computer and headphones and observed a 10-minute interactive Morse Code lesson on the computer. The pre-scripted lesson was situated in a virtual classroom consisting of a blackboard, a teacher agent, a seated avatar representing the participant, a seated colearner agent, and a shared desk (see Figure 1). The lesson consisted of a series of recorded "lectures," designed to teach the participant a total of nine Morse Code characters, interspersed with both transcription quizzes and short asides made by the colearner agent.

Experimental Design

Morse Code acquisition was chosen to simulate language learning, as language-learning software often uses animated agents. Because Morse Code conforms to a well defined set of rules [10], performance is easy to systematically evaluate by computer. In addition, using Morse Code for content matter (as opposed to, say, French) removed the need to quantify the effect that prior knowledge of other foreign languages (in this example, any other Romance language) might have on the participants' performance.

Stick figures were used to represent the teacher and colearner agents in order to minimize gender, age and race effects attributable to more realistic characters. In addition, a gender-neutral voice was used for the prerecorded statements to minimize the effect of gender [9]. Pre-tests confirmed that the voice did not have a discernible gender. Simple animated features, such as the teacher agent's talking mouth or blinking eyes, helped to make the environment more closely approximate typical interactive learning environments. In the virtual classroom, participants were identified as "Student A" (the participant's avatar) and "Student B" (the colearner agent) by names written on the chair backs.

During the lesson, the teacher agent administered six quizzes. Participants and the colearner agent were asked to transcribe sets of aurally transmitted Morse Code characters into text strings. For instance, after learning the Morse code for 'o', 's', 't' and 'e', the teacher played a series of beeps corresponding to the string "osteesnsaset." At the end of the transmission, the teacher asked either the participant or the colearner agent to provide their transcription. Letter by letter correction for the provided transcription was then placed on the virtual blackboard. Both the participant and the colearner had their transcriptions corrected on the board three times. The agent in the high-performing condition transcribed 75% of characters correctly during these quizzes and the low-performing agent had 33% of characters correct.

Correction of participant input was done on a character-by-character basis using a simple derivation of the string edit distance algorithm [10]. The algorithm provided scores and correction based on the most favorable interpretation of learner input. It aligned the maximal number of correct characters with the answer by dropping spurious characters and adding extra blanks, but without re-ordering characters.

To establish the socially supportive or competitive nature of the agent in each condition, the lectures and tests were punctuated by 16 pre-recorded comments made by the colearner. Two different sets of 16 comments each were delivered depending on whether the agent was socially supportive or competitive. The socially supportive agent made comments that were supportive or emphasized teamwork (e.g., “I like learning with you!”) The socially competitive agent’s comments were aggressive or compared performance (e.g., “I can beat that score!”).

After the lesson, participants were asked to complete a brief, web-based questionnaire. Some of the questions were of the form, “How well or poorly does each of the adjectives below describe [the agent/your feelings] followed by a ten-point Likert scale ranging from “Describes Very Poorly” (=1) to “Describes Very Well” (=10). The remainder of the questions were of the form, “How much do you agree or disagree with the following statements” followed by a ten-point Likert scale ranging from “Strongly Disagree” (=1) to “Strongly Agree” (=10).

The questionnaire served as an intervening task between the lesson and the final test. In the final Morse Code test, participants were shown the “virtual classroom” again and asked by the teacher agent to transcribe aurally transmitted characters. Six sets of 12 characters each were transmitted. No colearner agent was present during this test in any of the conditions. Correction and feedback were not provided between the six sets of characters.

Measures

Participant performance was based on number of errors made in the post-test. Results were systematically scored using the same character-by-character edit distance algorithm used to provide in-lesson correction.

The participant’s subjective response to the agent was measured by combining related statements and adjectives (some reverse-coded) into the following indices. *Socially competitiveness of the agent* (a=.73): socially competitive, socially supportive. *Intelligence of the agent* (a=.66): smart, trustworthy, unintelligent, “Student B performed better than I did. ” *Liking of the agent* (a=.87): fun, likable, distracting, frustrating, annoying. *Kindness of the agent* (a=.89): helpful, judgmental, caring, condescending, rude, and unkind. *Sociability of the user* (a=.65): polite and sociable. *Competitiveness of the user* was a single measure that indicates how competitive the user felt toward the agent. All indices were reliable.

RESULTS

Manipulation checks

The socially competitive agent was seen as significantly more competitive than the socially supportive agent (M=8.8, SD=1.6, vs. M=3.1, SD=1.4, $F(1, 40)=174.0$, $p<.001$) .

The high-performing agent was seen as significantly more intelligent than the low-performing agent (M=5.4, SD=1.2 vs. M=4.2, SD=1.5, $F(1,40)=18.0$, $p<.001$).

Performance measures

Participants working with high-performing colearners made significantly fewer errors on the final test than did participants with low-performing colearners, (M=17.0, SD=11.5 vs. M=27.9, SD=16.4, $F(1,40)=6.2$, $p < .02$). There was no effect for social competitiveness and no interaction.

Attitudinal measures

Participants with socially supportive agents liked the agent more than did those with socially competitive agents, (M=5.0, SD = 2.4, vs. M=3.1, SD=1.8, $F(1,40)=9.7$, $p<.001$, $h^2=.20$) and perceived the agent to be more kind (M=7.2, SD = 1.3 vs. M=3.1, SD=1.5, $h^2=.69$). Agent performance did not appear to affect the participants like or dislike of the agent or their perception of agent kindness.

There was a significant cross-over interaction with respect to the participant’s sociability toward the agent. Participants with socially supportive colearners felt more sociable toward the low-performing agent than toward the socially supportive high-performing agent, (M=6.4, SD=1.6, vs. M=5.3, SD=1.0), while participants with socially competitive agents felt more sociable toward the high-performing agent than toward the low-performing agent, (M=5.4, SD=1.2, vs. M=4.7, SD=0.8, $F(1,40)=6.15$, $p<.02$, $h^2=.13$). Socially supportive agent participants felt more sociable on the whole, than socially competitive agent participants, (M=5.9, SD = 1.4, vs. M=5.1, SD=1.07, $F(1,40)=4.95$, $p<.03$, $h^2=.11$). There was no significant effect for performance.

Socially competitive agent participants felt significantly more competitive toward the agent than did socially supportive agent participants, (M=5.5, SD=2.8 vs. M=3.8, SD=2.8, $F(1,40)=5.8$, $p<.05$). High-performing agent participants felt marginally more competitive toward the agent than did low-performing agent participants, (M=5.3, SD=2.9 vs. M=3.9, SD=2.8, $F(1,40)=3.6$, $p<.10$). There was no interaction.

DISCUSSION

Our key findings are that it is possible to make students perform better on average by creating interfaces with a high-performing colearning agent rather than a low-performing one, and that it is possible to make students like their colearning agent more on average by scripting its responses to be socially supportive rather than socially competitive. Complicating these findings somewhat is the indication that participants feel more sociable with a low-performing socially supportive agent than a high-performing socially supportive agent.

The larger implication of this finding that the *behavior* of the agent can affect individual's learning of the material. Even though participants were presented with the same material in all conditions, participants nonetheless learned more when their co-learner seemed to have a better grasp of the material.

One possible explanation for the positive effect a high-performing agent has on the participant's performance is that the colearner's performance provides a *benchmark* of how well a learner should do. This expectation could well influence a participant's level of learning effort. The primary competing explanation, that the participants are striving for *competitive performance,* is undermined somewhat by the fact that colearner competitiveness was not demonstrated to affect the performance, even though it did strongly affect the participant's subjective reactions to the agent.

FUTURE WORK

A few limitations of this work should be explored in future research. For this study, we ignored individual differences in participant preference for socially supportive, socially competitive, high and low-performing colearners. Although our results indicate a strong preference for socially supportive colearners on a whole, and a greatly improved performance to high-performing colearners overall, one useful follow-up to the current experiment would be to correlate preference and performance with participant performance on a Myers-Brigg type personality test [7]. It would also be of interest to perform additional tests where the colearning agent performed better or worse *relative* to the participant's performance.

Additional opportunities for further inquiry are experimental tests investigating the effects of multiple colearning agents, or no colearning agent at all. The structure of our particular experiment made it difficult draw direct comparisons to a no-agent condition, but ultimately it would be desirable to establish the degree and manner to which integrated learning environments with colearning agents could improve those without.

The current research represents the first systematic experiment using colearners who are not trying to be intelligent—in the machine learning sense. While intelligence likely provides many opportunities for successfully manipulating the performance and affect of students, this study suggests that a well-designed colearner can greatly influence performance and feelings.

ACKNOWLEDGMENTS

This research was supported by grants from the OMRON Corporation and the CHIMe Lab. We thank Scott Brave, Katherine Isbister, and Fred Turner for valuable comments and assistance during the course of this experiment.

REFERENCES

1. Azmitia, M. Peer interaction and problem solving: When are two heads better than one? *Child Development*, 59 (1988) 87-96.
2. Chan, T., Baskin, A. Learning Companion Systems. in *Intelligent Tutoring Systems*, ed. C. Frasson, and G. Gauthier. Ablex Publishing Corporation. Norwood, New Jersey, USA (1990)
3. Dillenbourg, P., Self, J. A Computational Approach to Socially Distributed Cognition. *European Journal of Psychology of Education*, 7 (1992) 353-372.
4. Johnson, D., Johnson, R., Skon, L. Student achievement on different types of tasks under cooperative, competitive, and individualistic conditions. *Contemporary Educational Psychology*, 4 (1979) 99-106.
5. Lester, J., Converse, S., Kahler, S., Barlow, S., Stone, B., Bhogal, R. The Persona Effect: Affective Impact of Animated Pedagogical Agents. *Proceedings of CHI 1997.* ACM Press (1997) 359 – 366.
6. Levitt, B., Nass, C. The lid on the garbage can: Institutional constraints on decision making in the textbook publishing industry. *Administrative Science Quarterly*, 34:2 (1989) 190 - 207.
7. Myers, I., McCauley, M., *A Guide to the Development and Use of the Myers-Briggs Type Indicator*, Consulting Psychologists Press, Palo Alto, CA (1992).
8. Piaget, J., *The language and thought of the child*. New York: Harcourt (1928)
9. Reeves, B., Nass, C. *The media equation: how people treat computers, television, and new media like real people and places*, Cambridge University Press, New York, NY (1996)
10. Taylor, D.W., "The Learning of Radiotelegraphic Code," *American Journal of Psychology*, LVI (1943) 319-353.
11. Vygotsky, L., *Mind in Society: The development of higher psychological processes*. Cambridge MA: Harvard University Press (1978)
12. Wood, E., Willoughby, T., Reilley, S., Elliot, S., DuCharme, M. Evaluating students' acquisition of factual material when studying independently or with a partner. *British Journal of Educational Psychology*, 65 (1995) 237-247.