All or Nothing: Levels of Sociability of a Pedagogical Software Agent and its Impact on Student Perceptions and Learning

Kathryn Hershey,

Punya Mishra,
Learning, Technology & Culture Program
College of Education,
Michigan State University

Ellen Altermatt
Psychology Department
Hanover College

To appear in *Journal of Educational Multimedia and Hypermedia*.
Please do not quote or cite without permission.

August 2004
Abstract

This paper reports the results of an experimental study on multimedia learning environments, which investigated the impact that increasing the social behaviors of a pedagogical agent has on students’ perceptions of social presence, their perceptions of the learning experience, and learning. Paradoxically, in this experiment students detected higher degrees of social presence in both the text only and the fully animated social agent conditions than students in the voice only and the static image of the agent with voice conditions. Furthermore, students had more positive perceptions of the learning experience in the text only condition. The results support the careful design of social behaviors for animated pedagogical agents if they are to be of educational value, otherwise, the use of agent technology can actually detract from the learning experience.
All or Nothing: Levels of Sociability of a Pedagogical Software Agent and its Impact on Student Perceptions and Learning

Recent research in technological learning environments has begun to focus on the educational benefits of including pedagogical software agents. The work of Reeves and Nass (1996) has demonstrated that in human-computer interactions humans will anthropomorphize the software even to the extent of applying social rules of human-human communication to the computer agents. The application of these rules to interactions with a software agent in an electronic learning environment may have educational benefits as well. Pedagogical software agents are animated interface agents in instructional environments that draw upon human-human social communication scripts by embodying observable human characteristics (such as the use of gesture and facial expressions).

Research in the past few years on the use of pedagogical software agents has expanded to cover a variety of pedagogical roles taken on by the agent. The Teachable Agent Group at Vanderbilt (TAG-V) has developed social agents who play the role of the “tutee” rather than the tutor. In this environment where students learn by teaching, students are able to adjust the agent’s attitude, and teach him or her relevant skills and concepts. The Tutoring Research Group at the University of Memphis, under Art Grasser, has developed a system called AutoTutor, which uses conversational agents that act as dialogue partners to assist learners in the active construction of knowledge. The Pedagogical Agent Learning Systems (PALS) laboratory at the University of Florida, under Amy Baylor has developed a multiple agent environment called Multiple Intelligent Mentors Instructing Collaboratively (MIMIC) in which the agents act as
mentors to students. Research on the MIMIC environment has focused on manipulating the characteristics of the agents such as programming the agents to offer feedback from an instructivist and constructivist theoretical perspective (Baylor, 2002). She has also investigated the impact of using multiple agents that serve as experts and/or motivators to students (Baylor & Ebbers, 2003). The agents in all of the aforementioned environments are responsive to user input.

However, even in a more didactic electronic teaching environment agents can maintain motivational and affective features. According to Atkinson (2002):

In particular, it may be possible to structure an example-based learning environment so that a lifelike character can exploit verbal (e.g., instructional explanations) as well as nonverbal forms of communication (e.g., gaze, gesture) within the examples themselves in an effort to promote a learner’s motivation toward the task and his or her cognitive engagement in it. (p. 416)

In traditional classroom settings “it is difficult to deny that teaching by its very nature involves some sort of intervention in the learning process of students in an attempt to facilitate their acquisition of desired educational outcomes” (Shuell, 1996). One form that this intervention takes is “providing cues as to which information in the material being studied is most important and the manner in which students might process the information” (p. 731). In a computer-based tutorial presentation gestures, facial movements, and tonal changes in voice exhibited by an agent can be used to indicate important and relevant information. Thus, current research on agents needs to be expanded to understand how and to what extent agent behaviors contribute to helping students learn from didactic lessons.
Preliminary research does suggest that lifelike agents can have a strong motivational effect (Lester, Converse, Stone, Kahler, & Barlow, 1997) and promote deeper cognitive engagement (Johnson, Rickel, & Lester, 2000; Mayer, Sobko, & Mautone, 2003). In a study by Moreno, Mayer, Spires, & Lester (2001) the inclusion of a pedagogical agent had a positive effect on interest and transfer.

According to what James Lester calls the persona effect (James Lester et al., 1997) the presence of a pedagogical software agent in a multimedia-learning module will have a positive effect on a learner’s motivation and engagement due to the range of social behaviors it can exhibit. In a recent study by Baylor and Ryu (2003), researchers found that the use of an animated pedagogical software agent was beneficial for participant’s perceptions of persona characteristics. Persona characteristics were described as the propensity of the agent to be engaging, person-like, credible, and instructor-like. However, animation was not always the single best way to positively impact students’ perceptions of persona characteristics.

There is also preliminary empirical evidence that there are constraints placed on educational benefits when designing multimedia environments. Research studies by Moreno and Mayer (1999) on the optimal design of multimedia environments have argued that certain design elements can have a positive impact on student learning. For instance, they argue that text and pictures need to be physically integrated and temporally synchronized (Mayer, 1997), and that auditory presentation of text was better than just presenting it on screen (Mayer, Heiser, & Lonn, 2001). Mayer, Heiser, and Lonn (2001) do, however, add a note of caution:
In our opinion, multimedia design principles should not be taken as blanket commandments but rather should be interpreted in light of theories of how people learn—such as the cognitive theory of multimedia learning (Mayer, 1997). For example, PowerPoint presentations—in which a presenter both speaks and presents words on screen can be effective even though words are presented in two modalities. (p. 196)

Pedagogical software agents represent a new paradigm for teaching and learning based on research in the areas of animated interface agents and interactive learning environments (Johnson et al., 2000). Furthermore, animated pedagogical agents have the potential to broaden the bandwidth of social communication between computers and students and increase student engagement and motivation (Johnson et al., 2000).

One important question that designers of animated pedagogical agents need to consider is just how much sociability they need to incorporate into the agent. Is having a voice and a picture sufficient to elicit the persona effect? Or do we need other social behaviors such as gestures, eye contact, movement on screen etc. As Reeves and Nass (1996) have argued, there are situations where people respond socially to computers (and other interactive media) even in the absence of programmed anthropomorphic behaviors. This is an important question because programming social behaviors is non-trivial and often, expensive. Clearly, pedagogical software agents with their greater range of “behaviors” such as gestures, intonation, eye-contact, movement etc. bring an additional layer of issues that research and design need to consider.

The study reported here examines the impact that increasing the social behaviors of a software agent has on students’ perceptions of their experience and learning. It was
hypothesized that the more lifelike the social behaviors are the more positive student perceptions will be of their experience (a persona effect) and the better their performance will be on a test of learning. The persona effect only states that the presence of an agent “can have a strong positive effect on student’s perception of their learning experience”. Thus we are also hypothesizing that by manipulating the expressivity and presence of the agent that students’ perception of the medium as a social entity will also become more salient.

In order to measure the salience of the medium as a social entity, we also looked at the degree of social presence perceived by the learners as they interacted and learned from multimedia tutorials that had different levels of social cues.

Measures of social presence can assist researchers in understanding the degree to which learners are perceiving an “illusion of nonmediation” in regards to their experience (Lombard & Ditton, 1997) and the salience of the communication medium (Short, Williams, & Christie, 1997). Social presence theory was developed by Short, Williams, and Christie (1976) to identify the impact of different communication media on the perceived satisfaction and efficiency of the communication. Social presence, as defined by Short, Williams, and Christie, is “the degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships…” (p.65). The focus from this perspective is on the communications media and its ability to impact the salience of interpersonal communications between individuals. However, in their review of research related to presence Lombard and Ditton (1997) noted that various studies have shown that humans will perceive the medium itself as a social actor, and/or perceive
a social entity within non-interactive mediums. In this study we are conceptualizing social presence as the salience of the perception of the medium as a social entity. By manipulating the presence and expressivity of an agent in a multimedia environment it is possible that human’s responses to psycho-social stimuli will create a more engaging experience that impacts their performance, their perceptions of the learning experience, and the salience of the medium as a social entity.

*Predictions*

We hypothesized that greater levels of social presence will be observed as the sociability of the pedagogical software agent is increased i.e. the agent exhibits greater degrees of social behavior (such as synchrony between lip movement and voice, gestures, intonation etc.). Furthermore, as the agent becomes more expressive and lifelike by exhibiting, human-like paralinguistic cues, students will have more positive perceptions of the learning environment and demonstrate a better recognition of facts presented in the tutorial.

Although we hypothesized that increasing the level of sociability of the software agent will increase perceptions of presence, produce more positive perceptions of the learning experience, and better performance, there are actually two different ways of thinking about how students will respond to the increased sociability of the agent across the four conditions. One argument is that increasing agent sociability (adding voice, voice plus picture, voice plus animated social agent) will lead to increased engagement, motivation and recall of information by virtue of the “persona” effect. Thus the order from least to most effective would be from condition to I to IV – as we increased the
social richness of each condition (from no agent to just voice, to voice plus image to finally the animated social agent).

On the other hand, it can be argued that adding further modalities to the plain text version could actually increase cognitive load (Sweller, 1999), preventing students from paying full attention to the information. Thus increased sociability could actually be harmful by distracting the student’s attention from the information to be learned. However, according to Mayer et al. (2003) social agency theory and cognitive load theory “…are not mutually exclusive, because both cognitive and social factors can contribute to how students learn from multimedia messages” (p. 424).

Methodology

In this experiment, students were asked to learn about nanotechnology through a multimedia learning module. This module was constructed using PowerPoint and Vox Proxy agent software. After they worked with the software students responded to a series of items measuring how they felt about the computer program and the extent to which they perceived it as a social entity (social presence), how they felt while viewing the program (perceptions of the learning experience), and their recognition of the information presented in the module (performance). These variables were developed based on existing research and extensive pilot testing.

Participants and Design

The participants in this study (n=116) were undergraduate students who were enrolled in a teacher education course. The experimental manipulation consisted of changing the levels of sociability of the agent. Every presentation for the four conditions included informational text and supplementary graphics.
Materials

The learning module was created using Microsoft’s PowerPoint software, in combination with Vox Proxy software. Vox Proxy is special software that provides an interface for users with no programming expertise to script actions for agents that can be incorporated into PowerPoint presentations.

The materials created to represent the levels of social richness in the agent’s behavior in the four experimental conditions were as follows:

Condition I: Text only

- This condition was very similar to what students would normally see in an online tutorial (see Figure 1).
- The tutorial consisted of informational text and graphics.

Figure 1: Text only condition

Condition II: Voice only

- In the voice only environment, the software read the text out in a monotone voice.

Condition III: Image of an agent plus voice

- This condition was identical to condition II except that a static image of the agent, Chuck, was included on the screen. This image, clearly that of a software agent,
not a real person, matched the “mechanical” voice that was reading out the text (see Figure 2).

Figure 2: Image plus voice

Condition IV: Social agent

• This condition included an animated social agent. The voice was modulated and less mechanical. The agent moved around on the screen, pointed at objects, and read through the text with synchronized lip movements.

• As in condition III the agent read the same text that was on the screen, however, there was the addition of a few comments to promote the perception of a social entity. The comments did not include any other details or information in order to avoid an interference effect in learning (Moreno & Mayer, 2000) (see Figure 3).
In terms of the design of the measurement instrument, the social presence measures consisted of bi-polar items using adjective pairs as parameters on a seven-point scale. The items were based on previously used measurement instruments designed to measure social presence in mediated environments (Lombard et al., 2000; Short, Williams, & Christie, 1976). Examples of the social presence adjective pairs include: personal-impersonal, unsociable-sociable, insensitive-sensitive, and warm-cold. The perceptions of the learning experience measures also used seven-point bi-polar adjective pairs. Some examples of the adjective pairs are: passive-active, bad-good, and excited-bored. The items for the performance measure were based on a 17 item multiple-choice test.

Procedure

The students were informed that they would be tested on their recall of the information and, thus, should pay close attention to the information presented. Students were randomly assigned to one of four conditions. Each condition held the information presented constant while manipulating the levels of sociability. The data were collected through a paper and pencil test and surveys because we did not want the results biased by a perceived presence of the computers (Nass, Moon, Morkes, Kim, & Fogg, 1997). The surveys were administered after the participants viewed the presentation. The entire study, including viewing the resource material, took approximately 45 minutes.

Scoring

Individual items that were highly correlated and measured perceptions of the learning experience were collapsed into a single scale. Items that measured social presence were also collapsed into a single scale. Cronbach’s alpha for the perceptions of
the learning experience scale was .91 and for the social presence scale Cronbach’s alpha was .94. In addition, we also computed the performance score of each participant by totaling the number of items answered correctly on the test.

Results

Significant effects were found for both social presence and perceptions of the learning experience. The means and standard deviations are listed in Table 1.

Table 1: Means and Standard Deviations for Measures of Perceptions of the Learning Experience, Social Presence, and Performance

<table>
<thead>
<tr>
<th>Group</th>
<th>Perceptions of the learning experience</th>
<th>Social presence</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text only</td>
<td>M = 4.9, SD = .7</td>
<td>M = 4.3, SD = .9</td>
<td>M = 12.7, SD = 3.2</td>
</tr>
<tr>
<td>Voice only</td>
<td>M = 4.4, SD = .9</td>
<td>M = 3.0, SD = 1.1</td>
<td>M = 12.4, SD = 2.0</td>
</tr>
<tr>
<td>Image of agent and voice</td>
<td>M = 4.3, SD = .9</td>
<td>M = 3.1, SD = .9</td>
<td>M = 12.3, SD = 2.6</td>
</tr>
<tr>
<td>Social agent</td>
<td>M = 4.7, SD = .7</td>
<td>M = 4.2, SD = 1.0</td>
<td>M = 13.3, SD = 2.4</td>
</tr>
</tbody>
</table>

aMaximum score = 7
bMaximum score = 17

Measures

Perceptions of the learning experience. Participants were asked to respond to a series of items designed to measure how they felt while viewing the learning module. Results from a one-way ANOVA indicate that there was a significant difference between conditions, F(3, 112) = 4.14, p = .008. Tukey post-hoc tests revealed that students in condition I (text only), (M = 4.9, SD = .7) had more positive perceptions of the learning experience than students in condition III (image plus voice), (M = 4.3, SD = .9). The difference was significant at a p=.01 level. There was not a statistically significant difference between the higher mean of condition I (text only) and condition II (voice
only), p=.08. Statistical significance was also not observed between the higher mean of condition IV (social agent) and condition II (image plus voice), p=.08.

**Social presence.** The students’ perceived higher degrees of social presence in condition I (text only) and condition IV (social agent). The results for the social presence measures were statistically significant, F(3, 112) = 14.40, p = .000. A Tukey post-hoc procedure was performed to detect differences between individual conditions.

There were statistically significant differences between condition I (text only), (M= 4.3, SD =.9) from both condition II, (M = 3.0, SD =.1.1) and condition III, (M = 3.1, SD =.9) (the voice only and the image plus voice conditions) at a p = .000 level. In addition, there was a statistically significant difference between condition IV, (M = 4.2, SD =.10) (social agent) from both condition II and III (the voice only, p = .000 and the image plus voice, p = .001).

**Performance.** There was not a significant effect of the experimental manipulations on performance, F(3, 112) = .909, p = .439. A possible explanation for this may be that there was a significant effect for gender in terms of performance F(1, 114) = 4.66, p =.03. Furthermore, due to random assignment males and females were not equally distributed among the conditions. There was not a significant effect for gender in terms of social presence F (1, 114) = .011, p = .917, or perceptions of the learning experience F(1, 114) = .508, p = .478.

**Discussion**

Paradoxically, it appears that the participants perceived higher degrees of social presence at both ends of the sociability continuum. For example, the text only and the fully social agent conditions were both seen as having a greater level of social presence
than the voice only and the image and voice conditions. A similar pattern was also observed with text only condition and the items that measured perceptions of the learning experience. The fact that the social agent would have a greater degree of presence is not surprising according to the persona effect. However, the fact that the text only condition shows the same pattern and has significantly higher perceptions of the learning experience is unexpected.

It may be possible that the mechanical characteristics of the voice only and image and voice conditions actually detracted from the social presence of the presentation. The findings of Baylor (2003)suggests that the inclusion of agent properties that are not life-like may actually detract from learning. Thus, the motionless image of the agent in condition III, and the mechanical voice in conditions II and III may actually interfere with learning. Baylor also states that an important attribute of agents is that they are perceived as “person-like” in order to establish a viable relationship with the learner. However, this does leave unanswered the question of why the complete lack of an agent led to high levels of presence.

Another contributing factor can be found in the work of Reeves and Nass (1996) (See also Mishra, Nicholson & Wojcikiewicz, 2001/2003). The evidence suggests that generating personality in a piece of software is not difficult and that users often attribute social characteristics to simple interfaces. This has been called Topffers Law (Mishra, et. al., 2001/2003) and states that almost all interfaces, however badly developed, have personality and that personality can appear through the subtlest of cues, text messages, layout, through the use of images and graphics. Of course the use of anthropomorphic agents, such as those used in this experiment will only enhance this effect.
So one can argue that, the participants in Condition I are behaving consistent with Topffer’s Law, However, the participants in condition II (mechanical voice) and condition III (static picture with mechanical voice) are suddenly offered additional cues that may not match their initial judgment of agency. Moreover, the social cues are non-lifelike (mechanical voice, static picture with a voice) which may actually detract from the perception of agency. It is only in condition IV (the fully sociable agent) where agency and personality can fully emerge.

Limitations and Future Directions

The results generated from this study provide a basis for further research into the use of pedagogical software agents in computer based tutorials. One limitation of this study was its inclusion of graphics, text and voice in the same condition. This may have produced a split-attention effect (Mayer & Moreno, 1998). There is also the possibility that a redundancy effect was produced (Mayer et al., 2001). However, if this is the case then we would have expected to see a redundancy effect in all three conditions which included a voice, which was not the case. Clearly, matters are not as straightforward as many proponents of agent based learning environments would argue. The fact that we found strong effects for the “least sociable” condition indicates that people’s perception of agency is complex and not easily captured. Thus, future research in this area should be expanded to look closer at both ends of the sociability continuum in computer-based tutorial environments.

Conclusion

Early empirical evidence is already beginning to demonstrate that the use of agents in multimedia learning environments is related to gains in learning outcomes
(Atkinson, 2002; James Lester et al., 1997). Although this study did not find a significant effect for measures of learning, our research did conclude that the text only and social agent conditions resulted in higher perceptions of social presence and the text only condition resulted in more positive perceptions of the learning experience. Clearly this is an arena ripe for further research. However, we can argue that the designers of multimedia tools need to think carefully about whether or not to include pedagogical agents. Agents that are not lifelike, and are more mechanical in their behaviors, may actually detract from the learning experience. Even fully expressive agents may improve performance in the short term but once the novelty effect has worn off they may become more of a nuisance to users, though this was not something that was studied in this experiment.

Users come into our classrooms with different learning styles and preferences. Our results with the text only and social agent conditions makes it clear that to accommodate the diversity of learners in our classrooms, we need to build flexibility into the design of multimedia environments. When agents are added to environments they should be lifelike and not mechanical in nature. Conditions such as a mechanical voice or an agent who does not exhibit social behaviors may actually detract from the learning experience. Clearly, generating lifelike pedagogical agents is a non-trivial task, requiring a significant investment of time and effort. The art of designing characters is a complicated one. Designers of educational software tools have to go beyond the purely cognitive aspects of working with computers and factor in the social and psychological aspects of character design as well. This makes our task far more challenging, bringing as it does the psychology and art of performance to educational technology design. As the
results of our experiment shows that in certain contexts it may actually be better to have no agent than a badly designed one. In other words, as in most things in life, either do it well or don’t do it at all.
Author Note

Kathryn Hershey Dirkin is a doctoral student in the Learning, Technology, and Culture program at Michigan State University. Punya Mishra is an assistant professor in the Counseling, Educational Psychology and Special Education department at Michigan State University. Ellen Altermatt is an assistant professor at Hanover College.

This research study was partially supported by funding from the Joe and Lucy Bates Byers Fellowship and an Intramural Research Grant Program at Michigan State University to the second author and a Michigan State University Summer Research Fellowship to the first author. We would also like to thank Dr. Matthew J. Koehler for his assistance with the data analysis. Correspondence concerning this article should be addressed to Kathryn Hershey 315 Erickson Hall, East Lansing, MI 48824.


