The Impact of Image and Voice of Pedagogical Agents

Yanghee Kim
Department of Educational Psychology and Learning Systems
Florida State University
United States
yyk3841@garnet.acns.fsu.edu

Amy L. Baylor, Ph. D
Director, Pedagogical Agent Learning Systems (PALS) Research Laboratory
Department of Educational Psychology and Learning Systems
Florida State University
United States
baylor@coe.mail.fsu.edu

Gabrielle Reed
Department of Computer Science
Florida State University
reed@cs.fsu.edu

Abstract: The purpose of this study was to examine the effects of image and voice of pedagogical agents on student perception and learning. Pedagogical agents were developed with differing image (expert-like vs. mentor-like) and voice (strong vs. calm vs. computer-generated), but with identical gesture, affect, comments, and gender. 109 undergraduates in a computer literacy course were randomly assigned to one of the six conditions. The results revealed a significant main effect for agent image on role perception: mentor-like image was perceived as more motivating, as hypothesized. Also, there was a significant main effect for voice: the strong voice was overall most motivating, while both human voices (strong or calm) were perceived as more affable, affective, credible, and facilitating learning than the computer-generated voice. There was no significant effect of agent image and voice on learning.

Introduction

Human and computer interaction is fundamentally social (Reeves & Nass, 1996). Researchers have designed human-like interfaces to promote social interaction between human and computer (Koda & Maes, 1996; Takeuchi & Naito, 1995; Walker, Sproull, & Subramani, 1994). In the studies, the users perceived computing as more useful, reliable, engaging, and interactive when they had a human face in the interface.

Pedagogical agents are life-like characters to facilitate learning in computer-based learning environments (Johnson, Rickel, & Lester, 2000). Researchers showed the positive impact of pedagogical agents for motivation and learning (Atkinson, 2002; Lester et al., 1997; Mulken, Andre, & Muller, 1998). More important, the presence of agent image significantly affected the learners' perception of pedagogical agents (Baylor & Ryu, 2003). When the image of an agent was presented, the learners perceived the agent as more credible. Also, voices are powerful indicators of social presence. People tended to assign age and gender to a computer according to the voices from the computer and behaved politely to the voices on the computer (Nass & Steuer, 1993; Walton & Orlikoff, 1994). When voices are different, the computer was perceived as a different social actor with a unique point of view.

However, very few studies examined more specific characteristics of image and voice of effective pedagogical agents. That is, to design a pedagogical agent as an effective learning partner, what types of image and voice will be more effective and appealing? Previous studies have shown that learners preferred to have a human voice than computer-generated one (Atkinson, 2002; Mayer, Sobko, & Mautone, 2003). Will learners be satisfied with any type of image and voice as long as they are human-like? We cannot manipulate the image and voice of human instructors, but

it will be worth delineating the image and voice of computerized instructors as more appealing. By providing a pedagogical agent with more amiable image and voice, we can effectively invite learners' attentions and facilitate their active engagement. Also, Baylor and Kim (2003) simulated pedagogical agents as an expert, a motivator, and a mentor in terms of four components of image, voice, animation, and affect. However, they did not clarify the relative contribution of each component to the role instantiation and student perception. This study aimed at identifying the type of agents' image and voice to be more desirable and appealing for undergraduates. The study examined the effect of the type of image and voice of pedagogical agents on role perception and learning in a controlled experiment.

Method

Agent Development

The pedagogical agents in the study were instantiated as simulated instructors and differentiated by image (expert-like vs. mentor-like) and voice (strong, calm, and computer-generated). The expert-like agent wore suits and looked stern; the mentor-like agent wore casual and looked comfortable. The strong voice was authoritative, assertive, and enthusiastic. The calm voice was soft, nice, and kind. Except for image and voice, the agents were controlled for gender, ethnicity, comments, gestures, and emotional expressions. Figure 1 presents the levels of the agents' image and voice in the study.

	Image	
	Expert-like	Mentor-like
		
Voice	Strong Calm Computer-generated	

Figure 1: Image and voice of agents

Participant

Participants were 109 undergraduates enrolled in an introductory computer literacy course in a large southeastern public university. 37.6% of the participants were male; 62.4% were female. The average age of the participants was 19.81 (SD=4).

Dependent Variable

The dependent variables were role perception, agent persona, and learning. Role perception included two sub-measures of agent role and general assessment. Agent persona was measured with four sub-measures of affect, affability, credibility, and facilitation of learning. For agent role, the participants answered the question, *What was the agent's primary purpose?* They selected either one or both the choices: 1) To be an expert and 2) To be a motivator. For general assessment of the agents, the participants assessed the agents in terms of two questions: 1) *How motivational the agent was?* 2) *How expert-like the agent was?* Agent affect was measured with three items: 1) *The agent was enthusiastic*, 2) *The agent was expressive*, and 3) *The emotion of the agent was natural*. Affability was measured with three items: 1) *I liked to learn from the agent*, 2) *The agent was enjoyable*, 3) *The agent was friendly*, and 4) *The agent was interesting*. Credibility was measured with four items: 1) *The agent was human-like*, 2) *The agent was instructor-like*, 3) *The agent was intelligent*, and 4) *The agent was knowledgeable*. Facilitation was measured with five items: 1) *The agent was motivating*, 2) *The agent helped me to concentrate on the information*, 3) *The agent improved my knowledge*, 4) *The agent made instruction interesting*, and 5) *The agent presented material effectively*. All the items were scaled ranging from 1 (Not at all) to 5 (Excellent). Learning was measured by an open-ended recall test, in which the students were asked to write down all the ideas they had been told by the agent (Mayer & Gallini, 1990). The number of the ideas in the students' answers were counted and coded by two raters. Inter-rater reliability was evaluated as Cohen's Kappa=0.90.

Procedure

Participants were randomly assigned to one of the six agent conditions. They were instructed to listen to and watch the agent. In the intervention, an agent was present and provided brief information how to plan instruction for about 10 minutes. The agents spoke without students' requests for more information, so all the students were exposed to the same amount of information. After that, the participants answered posttest questions regarding role perception and learning. The whole session took approximately twenty minutes.

Design and Analysis

The study used a 2 by 3 factorial design, in which agent image had two levels (expert-like vs. mentor-like) and agent voice had three levels (strong vs. calm vs. computer-generated). To analyze agent role, two-independent-sample χ^2 test was conducted. To analyze general assessment, multivariate analysis of variance (MANOVA) was conducted. For each sub-measure of agent persona (agent affect, affability, credibility, and facilitating learning) and learning, two-way analysis of variance (ANOVA) was conducted. Alpha was set as .05.

Results

Agent Role

The results revealed a significant main effect for agent image, χ^2 =21.2 p<0.05. The primary purpose of the expert-like image was rated as *To be an expert* (50%), *To be a motivator* (47.7%), and *To be both* (2.3%). The primary purpose of the mentor-like image was rated as *To be an expert* (27.7%), *To be a motivator* (69.2%), and *To be both* (3.1%). Agent voice did not affect the student perception of agent role. There was no interaction effect of agent image and voice.

General Assessment

The overall MANOVA results revealed a significant main effect for agent image, Wilks' Lambda=.926, F [2, 102]=4.06, p < .05. The mentor-like image (M=1.86, SD=0.91) was significantly more motivational than the expert-like image (M=2.20, SD=0.86). Also, the univariate result revealed a significant main effect for agent voice on the sub-measure, *How motivational the agent was*, F [1, 103]=3.30, p < .05. The agents with the strong voice (M=2.33, SD=0.84) was significantly more motivational than the agents with the calm voice (M=1.92, SD=0.79). There was no interaction effect between agent image and agent voice.

Agent Affect

There was a significant main effect for agent voice on agent affect, F [2, 103]=4.54, p < .05. Tukey HSD revealed that the agents with human voices (strong and calm) were significantly perceived as more affective than the agents with the computer-generated voice. There was no main effect for image on agent affect; there was no interaction effect between image and voice.

Affability

There was a significant main effect for agent voice on affability, F [2, 103]=4.99, p < .01. Tukey HSD revealed that the agents with human voices (strong and calm) were significantly perceived as more affable than the agents with the computer-generated voice. There was no main effect for image on agent affect; there was no interaction effect between image and voice.

Credibility

There was a significant main effect for agent voice on credibility, F [2, 103]=5.23, p < .01. Tukey HSD revealed that the agents with human voices (strong and calm) were significantly perceived as more credible than the agents with a computer-generated voice. There was no main effect for image on agent affect; there was no interaction effect between image and voice.

Facilitation of learning

There was a significant main effect for agent voice on agent affect, F [2, 103]=4.99, p < .01. Tukey post hoc revealed that the agents with human voices (strong and calm) were perceived as more facilitating learning than the agents with the computer-generated voice. There was no main effect for image on agent affect; there was no interaction

effect between image and voice.

Learning

There was no significant main and interaction effect of image and voice on the students' recall.

Discussion

The study aimed at examining the effects of the type of agent image and voice on role perception, agent persona, and learning. With regards to agent image, the results revealed that agent image significantly affected the students' perception of agent role. The perceived role of the mentor-like image was a motivator, whereas the perceived role of the expert-like image was an expert. In the similar fashion, the students assessed the mentor-like image as significantly more motivating than the expert-like image.

Regarding agent voice, the students perceived the strong voice as significantly more motivating than the calm voice. Consistent with previous studies, the results showed the students' preference for human voice over computer-generated voice. The students perceived both human voices (strong or calm) as significantly more affective, affable, credible, and facilitating their learning than the computer-generated voice.

Overall the results indicated that agent image more influenced the participants' perception of agent role, whereas agent voice more influenced their perception of agent persona. Among the six conditions, the mentor-like image with the strong voice was significantly perceived as more expressive, human-like, and motivating than the expert-like image with the computer-generated voice. The results suggest that the type of agent image and voice be significant predictors for students' motivation in pedagogical agent-based learning environments. Future research is expected to confirm the findings of the current study.

References

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology, 94*(2), 416-427.

Johnson, W. L., Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education, 11*, 47-78.

Koda, T., & Maes, P. (1996). *Agents with Faces: The Effect of Personification.* Paper presented at the Fifth IEEE International Workshop on Robot and Human Communication, Tsukuba, Japan.

Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhoga, R. S. (1997). *The persona effect: affective impact of animated pedagogical agents.* Paper presented at the CHI'97 Human Factors in Computing Systems., Atlanta, GA.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*(4), 715-726.

Mayer, R. E., Sobko, K., & Mautone, P. (2003). Social cues in multimedia learning: role of speaker's voice. *Journal of Educational Psychology, 95*(2), 419-425.

Mulken, S. V., Andre, E., & Muller, J. (1998). *The Persona Effect: How Substantial Is It?* Paper presented at the HCI-98, Berlin.

Nass, C., & Steuer, J. (1993). Anthropomorphism, agency, and thopoeia: Computers as social actors. *Human Communication Research, 19*(4), 504-527.

Reeves, B., & Nass, C. (1996). *The Media Equation: How people treat computers, television, and new media like real people and places*. Cambridge: Cambridge University Press.

Takeuchi, A., & Naito, T. (1995). *Situated Facial Displays: Towards Social Interaction*. Paper presented at the Conference of Human Factors in Computer System (CHI-95), Denver, CO.

Walker, J. H., Sproull, L., & Subramani, R. (1994, April 24-28, 1994). *Using a human face in an interface.* Paper presented at the Human Factors in Computing Systems, Boston, Massachusetts.

Walton, J. H., & Orlikoff, R. (1994). Speaker race identification from acoustic cues in the vocal signal. *Journal of Speech and Hearing Research, 37*, 738-745.

Acknowledgments

This work was sponsored by National Science Foundation Grant # IIS-0218692