Etiquette and Efficacy in Animated Pedagogical Agents: The Role of Stereotypes

Kristen N. Moreno[1] **Natalie K. Person**[2] **Amy B. Adcock**[1]

Richard N. Van Eck[1] **G. Tanner Jackson**[1] **Johanna C. Marineau**[1]

[1]University of Memphis, Memphis, TN [2]Rhodes College, Memphis, TN

{kmoreno | aadcock | rvaneck | gtjacksn | jmarinea}@memphis.edu, person@rhodes.edu

Introduction

Recent advances in the development of intelligent tutoring systems include the addition of animated conversational agents that play the role of tutors (Person and Graesser 2002). These pedagogical agents can be programmed to behave in accordance with many social expectations, such as displaying appropriate nonverbal behaviors (e.g., Cassell et al. 1994). Failure to conform to these interaction norms could be perceived by a human as a breach of etiquette. Indeed, Reeves and Nass (1996) present compelling evidence that humans apply the same rules of social interaction to computers as to other humans. For example, people are polite when interacting with computers, even though computers presumably lack the human capacity to appreciate courtesy.

The expectations based on rules of etiquette frequently facilitate social interactions among humans. It stands to reason, especially in light of the evidence presented by Reeves and Nass (1996), that this would be true for human-computer interactions as well. Perhaps adherence to rules of etiquette could even affect the usefulness of computer programs. For example, breaches of etiquette could distract the user from the task at hand or even cause discomfort or annoyance that would discourage the user from interacting with the program. The Tutoring Research Group at the University of Memphis has developed an intelligent tutoring system called AutoTutor (Graesser, Hu et al. 2001; Graesser, VanLehn et al. 2001), which features an animated agent. We have been keenly interested in discovering and understanding any characteristics that may affect the pedagogical efficacy of agents. The current study was the first in an ongoing line of research investigating these characteristics, with the ultimate goal of designing more effective, pleasant, and entertaining tutors.

Many expectations about people are based on stereotypes. A stereotype is a mental device that assists in simplifying and structuring an overwhelmingly complex social environment (Smith and Medin 1981). Social perceivers routinely categorize people on the basis of their social groups, such as ethnicity and sex (Allport 1954). The stereotypes associated with these social categories offer a wealth of information from which the perceiver may effortlessly infer expected traits and behaviors, thereby avoiding the practically insurmountable burden of systematically processing the plethora of information presented by each individual (e.g., Brewer 1988). Unfortunately, the information imparted by stereotypes is often inaccurate or inappropriate as a basis for social judgment. However, despite the problems inherent in reliance on stereotypes for information about individuals, people still routinely invoke stereotypes as a basis for expectations about interactions (see Macrae and Bodenhausen 2000), even if they do not consciously intend to do so (e.g., Gilbert and Hixon 1991).

In this research, we examine how people's stereotypic expectations about agents affect pedagogical efficacy. If a pedagogical agent were to fail to conform to stereotypic expectations, which would be a breach of etiquette in a sense, it could distract the learner and disrupt learning. On the other hand, if the agent were to behave too stereotypically, which would constitute a different (but related) breach of etiquette, this could potentially offend or disquiet the learner, which might also disrupt learning. The current research was concerned with whether people apply stereotypes of human social groups to computerized animated agents based on the agents' visually apparent demographic characteristics. If humans view computers and other media as social actors, as Reeves and Nass (1996) suggest, then it stands to reason that people may attribute stereotypic traits to computerized animated agents as they do to humans. Also, because we are ultimately interested in developing pedagogically effective animated agents for use in intelligent tutoring systems, we were interested in whether the visually apparent group memberships of pedagogical agents would affect learning in the tutor-learner interaction.

We first asked the participants in this study about their impression of an animated agent who was visually identifiable as a member of certain stereotyped social groups. The agent then delivered a tutorial on blood pressure. We assessed perceptions of the agent with regard to stereotypic traits and learning from the tutorial.

Method

Participants and Design

Thirty-nine participants (69% female) recruited from the Introductory Psychology participant pool at the University of Memphis received course credit in exchange for their participation. Ages ranged from 18 to 48 years with a median age of 20 years. Forty-nine percent of the participants indicated an African-American ethnicity, and 51% identified as Caucasian.[1] We randomly assigned these participants to one of four conditions defined by a 2 (ethnicity of agent: African-American or Caucasian) x 2 (sex of agent: male or female) between-subjects design. We also included participant ethnicity in the design.[2] Because participants rated the agent on four different stereotyping subscales, we analyzed the subscales as a repeated measure.

[1] Because there was only one observation each for the Asian, Native American, and Latino ethnic identities, we did not include these participants in the analyses.

[2] Because we assigned participants to conditions randomly, some cells *n*s were too small to include participant sex.

Stimulus Materials

We created the agents using Poser 4, a three-dimensional animation software product by Curious Labs, and compiled them using Microsoft Agent Character Editor. The manipulation of sex required the use of two different agents. To manipulate agent ethnicity we changed the skin, eye, and hair color; hairstyle; and shape of the lips, nose, and face. The voices were recorded live samples from four humans, males and females who spoke American English with a slight Southern accent or African-American Vernacular English (a dialect of Standard American English spoken primarily by African-Americans). The recordings contained exactly the same information, except for the name by which the agent introduced itself. We integrated the agents with the instructional content on blood pressure using Macromedia Authorware. The agents we created for this study appear in Figure 1. Although these agents are reproduced here in black and white, the participants saw them in full color.

Figure 1. Agents used in the study.

Procedure

Participants arrived at the laboratory in groups of 1 to 10 people. All instructions, stimuli, and materials were presented to each person on a desktop computer by Macromedia Authorware software, and participants wore headphones for the duration of the experiment. Participants first viewed the animated agent presenting some information about how to use the program to navigate through the experiment (e.g., using the back and forward buttons). Following this presentation, participants rated the agent on a series of stereotype-relevant traits using 7-point Likert-type scales. The agent then presented a tutorial on blood pressure, which participants viewed at their own pace. The tutorial was followed by 18 multiple-choice questions on blood pressure. Participants then completed a demographic survey and manipulation checks. Upon completion of the experiment, participants were fully debriefed and credited for their participation.

Results

Stereotyping Indices

We computed several indices to measure whether participants applied stereotypic expectations to the agents. Specifically, we averaged subsets of the trait ratings to form composite stereotyping indices for Caucasian males, African-American males, Caucasian females, and African-American females. The stereotypic index for each group contained traits that had been rated by pilot participants as highly characteristic of the stereotype for that group. All subscales had acceptable reliabilities, with Cronbach's αs ranging from .73 to .89. We analyzed the data using repeated measures analysis of variance (RM-ANOVA), which indicated that participants' responses differed across the stereotypic subscales ($F(3, 29) = 29.53$, $p < .001$). This main effect was qualified by an interaction with sex of the agent ($F(3, 29) = 4.00$, $p < .02$). Table 1 presents these results. Participants perceived little difference between the male and female agents on the two male stereotypic indices. However, they rated the female agents as being more stereotypically female than the male agents. In short, these data indicate limited evidence that participants applied stereotypes to these animated agents based on their visually identifiable sex (but not racial) group memberships. This parallels peoples' tendency to stereotype humans based on sex.

Table 1. Mean Stereotype Ratings by Index and Agent Sex.

	Agent Sex	
Index	Male	Female
African-American Male	3.75 (1.23)	3.52 (1.62)
Caucasian Male	4.15 (1.14)	4.38 (0.85)
African-American Female	2.75 (0.92)	3.43 (1.16)
Caucasian Female	3.03 (1.18)	3.62 (1.48)

Note. Numbers in parentheses are standard deviations.

Learning Index

We computed a learning index for each participant by scoring each test item as correct or incorrect and summing the number of correct responses. Scores ranged from 8 to 18, with a mean score of 14.5 (80.6%). An analysis of variance (ANOVA) computed on the learning index indicated that participants learned more from the male agents ($F(1, 31) = 4.42$, $p < .05$) and that Caucasian participants scored higher than did African-American participants ($F(1, 31) = 4.38$, $p < .05$). There were no significant effects of agent ethnicity. The means in Table 2 suggest that African-Americans underperformed only with the female agents, although this interaction is nonsignificant.

Table 2. Mean Proportions of Correct Answers on Learning Index by Participant Ethnicity and Agent Sex.

	Agent Sex	
Participant Ethnicity	Male	Female
African-American	.83 (.11)	.69 (.14)
Caucasian	.87 (.10)	.84 (.14)

Note. Numbers in parentheses are standard deviations.

Discussion

Reeves and Nass (1996) present detailed evidence that people apply human social interaction rules to computers and other media. In this research, we examined whether those social interaction rules extended to the application of stereotypes to animated agents based on visually apparent social group memberships. Based on these data, it appears that participants did apply social stereotypes based on sex, but did not apply stereotypes based on ethnicity, to the animated agents in this study who appeared to be members of stereotyped social categories.

Designing pedagogically effective animated agents to serve as tutors in intelligent automated tutoring systems is an important objective of the Tutoring Research Group (see Graesser, VanLehn et al. 2001). This line of research is intended to further this objective by determining the agent characteristics that facilitate effective communication and

learning between intelligent automated tutors and their students. In this case, we examined social categories as they relate to rules of etiquette that could affect the tutor-learner interaction. We have shown that sex of the animated agent that gives an anthropomorphic face to the intelligent tutoring system can affect learning. It may be that the female tutor broke with rules of etiquette about who should teach at a college level by not conforming to the stereotype of males as professors. Whether this effect is attributable to distraction or discomfort has not been resolved here, and remains an interesting question for future research. We have planned several studies for the near future to examine these issues more closely. In our next study, we will examine whether stereotypic expectations about domain knowledge interact with the demographic characteristics of agents to affect learning. For example, do people learn more effectively about car repair from an agent named Joe who wears greasy overalls, or can they learn just as effectively about this topic from an agent named Nancy in a pink apron?

This study was the first in a line of research intended to further understanding of human-computer pedagogical interactions by determining the situations and characteristics that facilitate effective communication and learning between computer-generated tutors and their students. We are very interested in the role that stereotypic information plays in facilitating or inhibiting learning from animated agents. Current educational trends suggest that computerized instruction will play an increasingly larger role in the education of many students, as evidenced, for example, by the increased availability of computers in educational environments and the growing number of web-based courses. As technology continues to be integrated into instruction, it becomes increasingly important to understand the factors that affect the way people learn from computers. An understanding of the etiquette of human-computer interaction could potentially yield great insight into this problem.

Acknowledgments

This research was supported by grants from the National Science Foundation (REC 0106965) and the Department of Defense Multidisciplinary University Research Initiative (MURI) administered by the Office of Naval Research under grant N00014-00-1-0600. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of ONR or NSF. The authors thank Art Graesser for insightful comments on an earlier draft of this paper and Ben Binford, Brook Bissinger, and Nicole Pera for assistance in data collection.

References

Allport, G. W. 1954. *The Nature of Prejudice.* Reading, MA: Addison-Wesley.

Brewer, M. B. 1988. A Dual Process Model of Impression Formation. In R. S. Wyer Jr. and T. K. Srull eds., *Advances in Social Cognition* 1:1-36. Hillsdale, NJ: Erlbaum.

Cassell, J., Pelachaud, C., Badler, N., Steedman, M., Achorn, B., Becket, T., Douville, B., Prevost, S., and Stone, M. 1994. Animated Conversation: Rule-Based Generation of Facial Expression, Gesture and Spoken Intonation for Multiple Conversational Agents. *Computational Graphics* 28:413-420.

Gilbert, D. T., and Hixon, J. G. 1991. The Trouble of Thinking: Activation and Application of Stereotypic Beliefs. *Journal of Personality and Social Psychology* 60:509-517.

Graesser, A. C., Hu, X., Susarla, S., Harter, D., Person, N. K., Louwerse, M., Olde, B., and the Tutoring Research Group 2001. AutoTutor: An Intelligent Tutor and Conversational Tutoring Scaffold. *Proceedings for the 10th International Conference of Artificial Intelligence in Education*. San Antonio, TX.

Graesser, A. C., VanLehn, K., Rosé, C. P., Jordan, P. W., and Harter, D. 2001. Intelligent Tutoring Systems with Conversational Dialogue. *AI Magazine* 22(4): 39-51.

Macrae, C. N., and Bodenhausen, G. V. 2000. Social Cognition: Thinking Categorically about Others. *Annual Review of Psychology* 51:93-120.

Person, N. K., and Graesser, A.C. 2002. Pedagogical Agents and Tutors. In J.W. Guthrie ed., *Encyclopedia of Education*. New York: Macmillan. Forthcoming.

Reeves, B., and Nass, C. 1996. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places.* Cambridge: Cambridge University Press.

Smith, E. E., and Medin, D. L. 1981. *Categories and Concepts.* Cambridge, MA: Harvard University Press.