Mental Models of Robotic Assistants

Sara Kiesler, Jennifer Goetz
Human Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA 15213
412 268-5919
kiesler@cs.cmu.edu; jenga@andrew.cmu.edu

ABSTRACT
If robotic assistants are to be successful, people will need appropriate mental models of what these robots can do and how they operate. We are developing techniques for measuring people's mental models of interactive robots and social agents. We aim to measure the content of these models—if they are anthropomorphic or mechanistic—and the richness of these models (how elaborate or sparse they are; how much confidence people have in them). We report progress here.

Keywords
Social interfaces, mental models, anthropomorphism.

THEORETICAL FRAMEWORK
Mental models refers to people's conceptual frameworks to support their predictions and coordination in the world. Researchers working on believable agents and emotional robots have used anthropomorphism in their designs, e.g., [1], but measures of these processes are undeveloped.

Content of mental models
People can integrate ostensibly incompatible images and categories into a consistent, anthropomorphic mental model [3]. A joking robot could evoke the concepts, *machine* and *humorous person*, leading to the concept, *cheerful robot*, which incorporates mechanical and anthropomorphic features.

Richness of mental models
Research shows that people have a sparse, simple mental model of those with whom they have little direct experience [2]. With experience, people's mental models become richer and more complex. A first step in understanding this process is to devise measures of it.

EXPERIMENTS
We conducted 2 experiments with an interactive robot to develop measures of the content and richness of people's mental models of robots. We tested the validity of our measures by examining how people's scores varied on the scales when they interacted with robots that varied in appearance or behavior. If a measure of anthropomorphism is valid, then people should give higher scores when they interact with a humanoid robot than with a machine-like robot. If a measure of mental model richness is valid, then it people should give more rating and more complex ratings as they gain experience with the robot [2, 5].

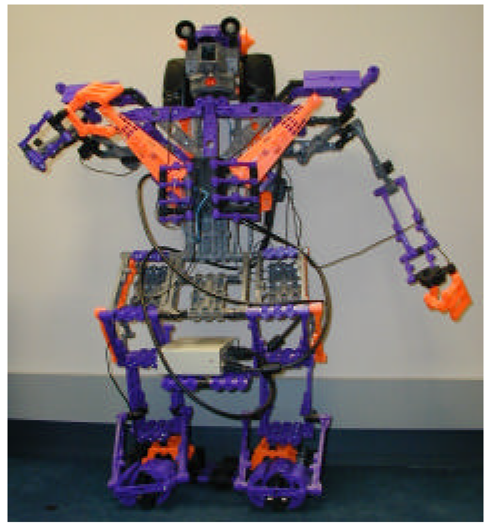

Figure 1. Toy Robot-Man.

Measures
Content.
To measure anthropomorphism, we adapted existing scales of sociability, intellect, and personality in people [4, 6]. We also developed a measure of mechanistic mental models.

Richness. Richness of a mental model is correlated with confidence in making ratings [2], reflected in the number of trait ratings they are willing to make [5].

Robots
We constructed two kit robots with small speakers and a wireless connection to a remote laptop. We used a quasi-Wizard of Oz procedure. One robot was constructed as a man about 3 ft. high; the other was constructed as a vehicle, about 3 ft. long. The experimenter controlled the robot's speech through a predefined script on the laptop.

Experiment 1: Toy Robot-Vehicle vs. Toy Robot-Man
In this experiment, college student volunteers conversed with the toy robot-vehicle (n = 10) or robot-man (n = 11) and answered interview questions from the robot. In this study we used scales for social and intellectual evaluation and ratings of human likeness to assess the content of the mental model, and the number of trait ratings people were willing to make to assess the richness of the mental model.

Results
The appearance of the robot did not affect participants' anthropomorphism, suggesting that the sociability and intellect scales did not measure anthropomorphism sensitively. However, participants made more trait ratings

when rating obvious and simple traits of the robot (e.g., cheerful) than when rating more complex, less observable

Figure 1. Toy Robot-Vehicle.

traits (e.g., responsible). This difference, 87% vs. 63% of the items offered for rating, was highly significant ($p < .01$) and consistent with the literature on out-groups and sparse mental models [5]. The results suggested that participants had a measurably sparse rather than rich mental model. We needed a better measure of anthropomorphism.

Experiment 2: Robot With/Without a Visible "Computer"

Twenty students and staff played a simplified 5-item desert survival decision-making game with the robot shown in Fig.1. In one condition of the experiment, we attached visible hardware—an external modem with cables—to the robot and in the other condition we presented the robot without the hardware. The purpose of this manipulation was to test whether the addition of a visible "computer" would alter the content or richness of the mental model.

Measures

We added a measure of anthropomorphism, adapting the Big Five [4]. We also devised a measure of mechanistic mental models, using these ratings: complex, obsolete, intuitive, works quickly, usable, durable, powerful, reliable, accurate. Three factors accounted for 67% of the variance. The items complex, (not) obsolete, quick, intuitive, and usable loaded on the scale, "advanced." The items, reliable and accurate loaded on the scale, "reliability." The items durable and powerful loaded on the scale, "power."

Results

Richness. Participants rated 100% of the mechanistic traits indicating they had a rich mental model of the robot as a machine. They were less willing to rate all the Big Five traits (see Figure 2). As in experiment 1, participants were most likely to rate the robot on extraversion—93% of the time. They were least likely to rate the neuroticism and openness to experience items—68%.

Content. The added hardware caused participants to have a less positive perception of the robot's reliability ($p = .06$), a more positive perception of its power ($p < .05$). At the same time, participants in the hardware condition had a slightly less positive perception of the robot on the anthropomorphic (Big Five) items.

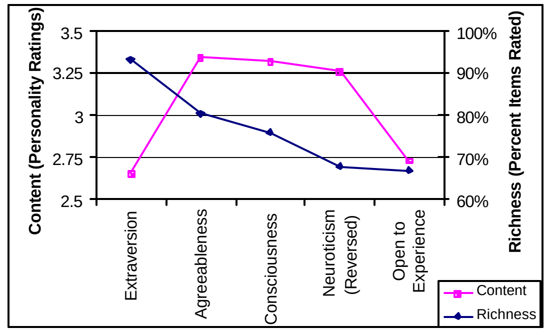

Figure 2. Mental model content and richness

Our results suggest progress in developing measurements of anthropomorphism and mechanistic mental models of robots, and independent measures of the richness of these models. Our measures now differentiate people's responses to different robots and also have face validity.

FUTURE RESEARCH

Our future work aims to develop reliable and valid measures of people's responses to robots, social agents, and robotic assistants. We are currently engaged in a study to measure confidence (richness) automatically using response times and extreme ratings. We believe this work will aid assessment of people's trust in interactive robots, and the degree to which they have appropriate conceptions of them.

ACKNOWLEDGMENTS

We acknowledge NSF grants # 0085796 & #1120180. We thank Allyson Pottmeyer, Sebastian Thrun, Nick Roy, Joelle Pineau, Jamie Schulte, Mike Montemerlo, Greg Baltus, and Greg Armstrong.

REFERENCES

1. Breazeal, C. (2000). Sociable machines: Expressive social exchange between humans and robots. Doctoral dissertation, Department of Electrical Engineering and Computer Science. MIT.
2. Gill, M. J., Swann, W. B., & Silvera, D. H. (1998). On the genesis of confidence. *Journal of Personality and Social Psychology*, *75*, 1101-1114.
3. Hintzman, D. (1986). Scheme abstraction in a multiple-trace memory model. *Psychological Review*, *93*, 441-428.
4. John, O. P., Donahue, E. M., & Kentle, R. (1991). The Big-Five Inventory. Tech. Report, IPSR, Univ. of California, Berkeley, CA.
5. Linville, P. W., Fischer, G. W., & Salovey, P. (1989). Perceived distributions of the characteristics of in-group and out-group members: Empirical evidence and a computer simulation. *Journal of Personality and Social Psychology*, *57*, 165-188.
6. Warner, R.M. & Sugarman, D. B. (1986). Attributions of personality based on physical appearance, speech, and handwriting. *Journal of Personality and Social Psychology*, *50*, 792-79.