in 6th ACM/IEEE International Conference on Human-Robot Interaction,
March 6-9, 2011, Lausanne, Switzerland.

Collaboration with an Autonomous Humanoid Robot: A Little Gesture Goes a Long Way

Kevin O'Brien, Joel Sutherland, Charles Rich and Candace L. Sidner
Computer Science Department
Worcester Polytechnic Institute
Worcester, MA 01609
(kobrien | jdsuth | rich | sidner)@wpi.edu

ABSTRACT

We report on an experiment in which a human collaborates with a small, autonomous, humanoid robotic toy. The experiment demonstrates that the robot's use of two simple gestures, namely orienting its head toward the addressee when it speaks and raising its arm in the direction of objects it refers to, significantly improve the human's perception of the robot's interaction skills and quality as a collaborator.

Categories and Subject Descriptors

H.1.2 [**Information Systems**]: User/Machine Systems

General Terms

Human Factors

1. EXPERIMENTAL SETUP

Figure 1: Robosapiens RSMedia robot.

The robot used in our experiment is a WowWee Robosapiens RSMedia (see Figure 1), which we named "Larry." It has a total of 11 degrees of freedom, including rotating its head from side to side and up and down, and one degree of rotation at each shoulder. Its eyes are illuminated, but do not move. The shoulder joints are canted so that when the arm rotates, it also moves away from the robot's body, which allows it to roughly gesture toward objects in the surrounding environment.

The robot can walk forward, backward and turn. In this experiment, we restricted its motion to a straight line by placing a wooden guide on the floor between its feet (see Figure 2). The robot has a speaker built into its body for playing prerecorded voice segments.

The robot runs Linux and comes with a proprietary behavior editor, which we used to program it. (Unfortunately, the RSMedia system software was not very reliable–see below). The robot can also receive infrared control signals, usually from a hand-held controller. In our experiment, however, we generated control signals from an infrared transmitter (see USB-UIRT in Figure 2) connected to a Windows XP computer. The human participant spoke to the robot using a head-mounted microphone connected to the same computer, which interpreted the speech using the Microsoft speech recognizer and used Girder home automation software to generate the appropriate infrared control signals.

HRI'11, March 6–9, 2011, Lausanne, Switzerland.
ACM 978-1-4503-0561-7/11/03.

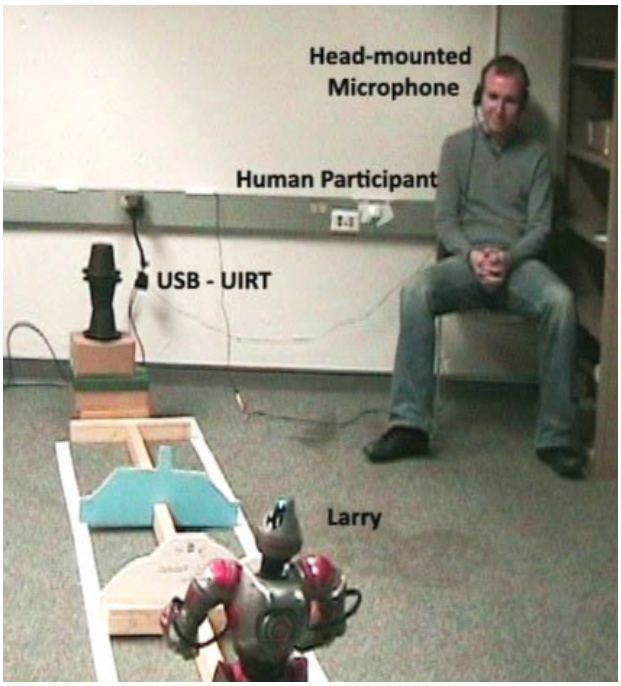

Figure 2: Experimental setup.

2. STUDY DESIGN

We designed a two-armed, between-subjects study. The two conditions differ only in whether or not the robot:

- orients its head toward the human while it is speaking,
- and raises its arm in the direction of objects referred to.

In the *gesture* condition, it performed both of these gestures; in the *control* condition, the robot's head remains in the level forward position and its arms at its side throughout the interaction.

In order to make the human-robot interaction collaborative, we created a science-fiction themed story in which the participant and the robot need to work together to overcome obstacles and plant a bomb to destroy an alien enemy fortress (the black tower at the back of Figure 2). Using prerecorded voice segments, the robot communicated the

story and asked the human to collaborate by retrieving and using various objects. For example, the robot says "Use the scissors to cut the power cable to the generator" (and in the gesture condition, extends its arm toward the desk on the left side of the room where the scissors are lying). The participants were instructed that the robot only understands three words: "affirmative," "continue" and "help." Participants were videotaped and asked to fill out an online questionnaire after completing the interaction.

We evaluated four subjective and one objective hypotheses. The following three subjective hypotheses were each evaluated using 7-point Likert scale questions (see examples below). Compared to the control condition, we hypothesized that in the gesture condition participants will:

(1) think the robot has better interaction skills (14 questions, e.g., "Larry responded well to me.")

(2) think the robot is a better collaborator (6 questions, e.g., "Larry and I worked well together.")

(3) enjoy the interaction more (5 questions, e.g., "I felt comfortable during the task.")

We also asked each participant how many minutes they thought it took to achieve the interaction goal, and hypothesized that in the gesture condition participants will:

(4) perceive that completing the interaction took less time.

Finally, we measured the actual time taken to achieve the interaction goal, and hypothesized that in the gesture condition participants will:

(5) complete the interaction in less time.

3. RESULTS

	Gesture	Control
male	8	12
female	5	1

A total of 36 participants between ages 18 and 52 were recruited from the WPI community. 10 participants were eliminated due to system software failures which prevented the interaction from being completed. The gender distribution of the remaining 26 participants is shown at the left.

Hyp.	P-Value
(1)	0.001 **
(2)	0.057 *
(3)	0.71
(4)	0.26
(5)	0.84

For the first four, questionnaire-based hypotheses, we used a grouped Mann-Whitney U Test resulting in the p-values shown at the left. Notice that the better interaction skills (1) and collaborator (2) hypotheses are very strongly and strongly supported, respectively. There was no significant difference in enjoyment (3) between conditions. The average interaction time (about 6 minutes) did not differ significantly (5) between conditions; although participants perception of the interaction time (4) tended to be shorter in the gesture condition.

4. DISCUSSION

The results above are consistent with a large literature on the important role that gestures, especially gaze [1, 3] and pointing [2], play in human communication and collaboration. Many other human-robot interaction researchers, e.g., [4, 5], have also implemented and evaluated head and arm gestures for robots. This experiment is valuable as a very simple and direct illustration of these effects.

Figure 3: Control condition example.

In addition to the quantitative results above, we would also like to share some anecdotal observations of the participants in the two conditions. Notice in Figure 2, which was taken in the gesture condition, that the human participant is sitting straight up, looking at the robot (whose head is turned towards the human), and generally looks engaged in the interaction. In comparison, the control condition participant shown in Figure 3 is looking at the floor in a distracted, unengaged fashion (the object on the floor is not a current topic of discussion with the robot). These differences were typical of the interactions in both conditions.

Finally, it is worth mentioning the possibility of a gender confound due to the uneven distribution of males and females between the two conditions.

5. ACKNOWLEDGEMENTS

This work is supported in part by the National Science Foundation under awards IIS-0811942 and IIS-1012083.

6. REFERENCES

[1] M. Argyle and M. Cook. *Gaze and Mutual Gaze*. Cambridge University Press, New York, 1976.

[2] C. Goodwin. Gestures as a resource for the organization of mutual attention. *Semiotica*, 62(1/2):29–49, 1986.

[3] A. Kendon. Some functions of gaze direction in two person interaction. *Acta Psychologica*, 26:22–63, 1967.

[4] B. Mutlu, T. Shiwa, H. Ishiguro, and N. Hagita. Footing in human-robot conversations: How robots might shape participant roles using gaze cues. In *Proc. 4th ACM/IEEE Int. Conf. on Human-Robot Interaction*, pages 61–68, 2009.

[5] C. L. Sidner, C. Lee, C. Kidd, N. Lesh, and C. Rich. Explorations in engagement for humans and robots. *Artificial Intelligence*, 166(1-2):104–164, 2005.