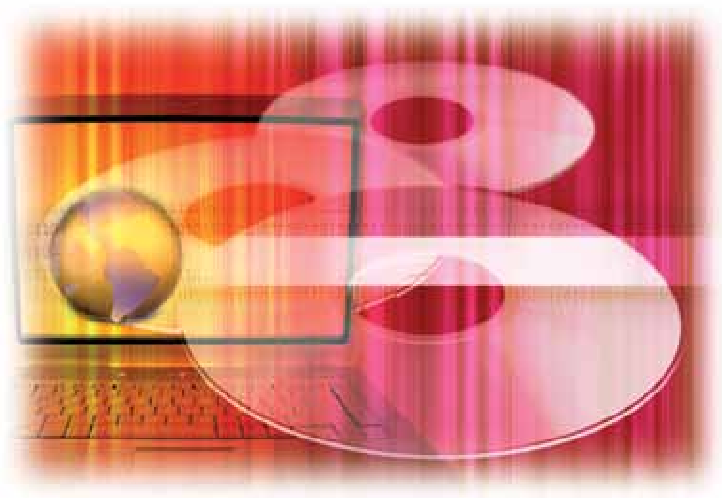
Playdates with Robots

Victor Ng-Thow-Hing
Honda Research Institute USA

Sandra Okita
Teachers College, Columbia University



Studies of different robot features and behaviors are leading to the development of models for more effective peer-based interaction between robots and children.

Although depictions of robots in films have created unrealistic expectations of how real robots should behave, the ongoing evolution of the physical capabilities continues to surprise and astound us. For example, humanoid robots can now run and open containers with their hands.

In these and similar tasks, robots interact with fairly predictable elements of the environment, like the ground or a jar. In contrast, when interacting with humans, robots must deal with a potentially large variety of responses.

Figure 1. A robot and child playing a table-setting game.

Underscoring the importance of solving the problems this presents, in 2011, the US government allocated $70 million to its National Robotics Initiative to develop robots that work cooperatively with humans. In fact, the goals of the area of research known as *social robotics* are to understand and successfully create effective interaction between robots and their human partners.

LEARNING WITH KIDS

We started the Learning with Kids project to investigate the use of robots as partners in a novel form of peer-based education (http://hri.willowgarage.com/workshops/RSS2011/downloads/NgThowHing.pdf). A humanoid robot has familiar body parts, such as a head and arms, to which young children can relate. A robot offers intimate, one-on-one interaction combined with untiring patience and the capability to record and personalize lessons.

By reviewing or explaining concepts using a robot that acts as a peer rather than an authoritative teacher, it's possible to engage children in learning in a new and exciting way. Children can use speech rather than typing on a keyboard to communicate with a humanoid robot. They look toward its "head" when addressing it or awaiting a response.

Our interdisciplinary collaboration between roboticists at the Honda Research Institute USA and a cognitive scientist from Columbia University made it possible to combine technically advanced robotic systems with carefully designed experiments and observation of human behavior. Initially, we had to make many decisions such as whether the interaction content was more important than the timing of feedback for creating engaging interaction. We conducted a series of experiments that sought to examine a wide range of robot behavior and its effect on children's behavior and level of participation in interaction with the robot.

Our goal was to develop tools that help children and robots engage in learning activities and to identify features and conditions that elicit active social interaction. Children ages four to 12 years volunteered to participate in our study by spending time engaging in game play, as shown in Figure 1, and storytelling activities with a humanoid robot.

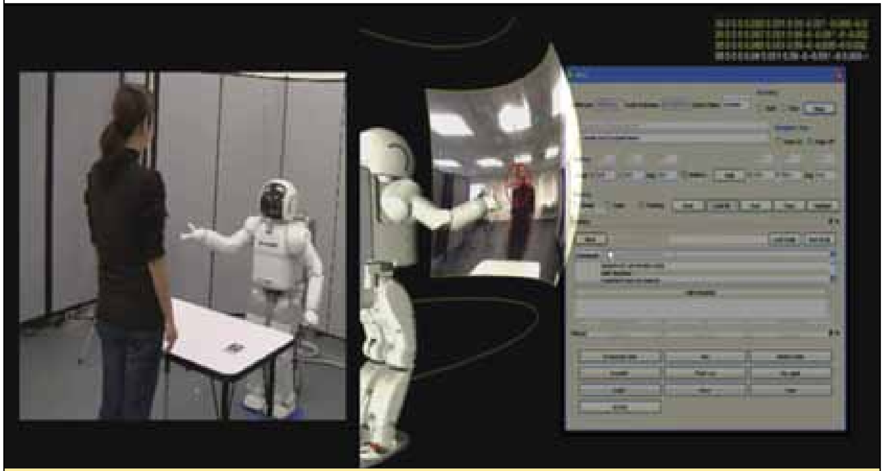


Figure 2. **The Wizard of Oz tool (right) controls the interaction (left). The hidden WoZ operator can click directly on the video image to direct the robot's gaze or pointing arm.**

Focusing on children was both challenging and rewarding. Although they generally have a shorter attention span than adults, we found that young children assign mixed attributes of living creatures and machines to robots, resulting in fewer preconceptions of what robots are capable of doing (S.Y. Okita and D.L. Schwartz, "Young Children's Understanding of Animacy and Entertainment Robots," *Int'l J. Humanoid Robotics*, vol. 3, no. 3, 2006, pp. 393-412). This openness makes children very interesting, unbiased subjects when conducting human-robot interaction studies.

WIZARD OF OZ

The problem with creating fluid, reliable robot interaction is that robot perception requires a high degree of accuracy and robustness. Generating expressive activity such as body motions or speech is within the robot's full control. However, its human partner's speech or nonverbal body language is subject to misinterpretation, which can trigger a chain reaction of incorrect behavior responses from the robot. This can quickly sabotage an experiment.

We used the Wizard of Oz (WoZ) technique to robustly prototype scenarios. Like the wizard in its namesake film, a hidden operator remotely monitors and controls the robot's verbal and nonverbal behavior. As Figure 2 shows, the operator can click directly on the camera view from the robot's head to determine where to make the robot look or point its arm. The operator can also select from a variety of pregenerated speech utterances or partake in freeform conversation.

Using WoZ, we could prototype with relatively little effort to determine the benefits of recognizing and responding to particular perceptual cues. This focused our research efforts on promising technological targets instead of wasting time using methods that might have a negligible effect.

WoZ also had an important role in benchmarking our developed techniques. To compare the performance of our automated interaction models against robots controlled by skilled human operators, we needed to develop measures that enabled us to quantify differences in the children's behaviors in each scenario.

MEASURING ENGAGEMENT

We used experimental and ethnographic methods to evaluate differences in the children's interactions with robots under different conditions. We measured interactions using multiple modalities—audio, video, skin conductance—and viewpoints.

We examined the audio and video recordings for social cues in the human partners' behaviors—such as where they directed their gaze and their utterances and body posture changes. We used skin conductance sensors to measure physiological arousal in the children. By counting the number of occurrences of these events over different time intervals under different conditions, we could quantify the effects of different experimental conditions.

Interaction styles

In one study, we examined whether high-level interaction styles such as lecturing, cooperative activities, or parallel play and familiar play routines such as turn-taking scenarios or setting the dinner table influence how children interact with robots.

The study revealed that children who interacted with the cooperative robot were more engaged and learned more from the interaction compared to using the lecture style or parallel play. The younger children performed as well as the older children in recall tests after interaction such as the table-setting game shown in Figure 1.

We noticed that a child's ability to pretend or engage in collaborative play was often constrained by what the robot could do in response—for example, timing and limitation of response. We found that, when the robot mentioned familiar social references such as going to the zoo or telling a familiar story, the child sustained a social relationship with the robot (S.Y. Okita, V. Ng-Thow-Hing, and R.K. Sarvadevabhatla, "Multimodal Approach to Affective Human-Robot Interaction Design with Children," *ACM Trans. Interactive Intelligent Systems*, vol. 1, no. 5, 2011, pp. 1-29). This suggests that, until robots have the intelligence to respond to any dynamic situation, they should follow a well-known script or scenario to guide the interaction.

Manipulating behaviors

We also conducted a series of studies manipulating different low-level behaviors such as using a human voice versus a robot-like voice, attention level (gaze response when the child speaks), and speech-synchronized gestures (speed and degree of expression) to see how specific features elicit responses from humans, including eye contact, number of responses, and length of engagement.

The robot's voice seemed to matter more for children 4 to 5 years of age, as their eye contact and engagement level dropped in conversation when the robot used a monotone voice but was sustained longer with a human-like voice. Having the robot respond to a child's speech using nonverbal attentional gaze was just as effective as verbal acknowledgments for short interactions. However, for longer interactions, verbal dialog became more important.

We created speech-synchronized gesture models that, depending on how fast we played the body motions, gave positive impressions such as being excited or confident at increased speeds and negative impressions such as being nervous or sad at slower speeds.

Game scenarios

If a robot can persuade a child to stand closer, it can perceive more details about the child. Therefore, in our third line of work, we examined whether using familiar game scenarios such as playing "Captain, May I?" could influence proxemics—physical distance—between humans and robots (S.Y. Okita, V. Ng-Thow-Hing, and R.K. Sarvadevabhatla, "Captain May I? Proxemics Study Examining Factors that Influence Distance between Humanoid Robots, Children, and Adults during Human-Robot Interaction," *Proc. 7th ACM/IEEE Int'l Conf. Human-Robot Interaction* [HRI 12], ACM, 2012, pp. 203-204).

We found that the robot discovered each child's initial comfort zone, then we explored whether additional verbal and nonverbal prompting such as beckoning could reduce that physical distance even further. Both verbal and nonverbal prompting were effective, but interestingly, when the robot asked permission, children allowed it to come closer compared to when they heard the robot announce its actions or saw the robot step forward without any warning.

Based on our studies of different robot features and behaviors, we're beginning to develop models for more effective peer-based interaction between robots and children. Being able to estimate and predict behavioral responses in their human partners will enable robots to dynamically modify their style of communication while attempting to fulfill their goals. Eventually, such models will lead to more pleasant and engaging social experiences while reducing frustration and boredom for children having playdates with robots.

Victor Ng-Thow-Hing *is a principal scientist at the Honda Research Institute USA, Mountain View, California. Contact him at vng@honda-ri.com.*

Sandra Y. Okita *is an assistant professor of technology and education at Teachers College, Columbia University. Contact her at okita@tc.columbia.edu.*

Editor: Kelvin Sung, Computing and Software Systems, University of Washington, Bothell; ksung@u.washington.edu