HUMAN-COMPUTER INTERACTION, 2004, Volume 19, pp. 61–84

Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial

Takayuki Kanda, Takayuki Hirano,
and **Daniel Eaton**
ATR Intelligent Robotics and Communication Laboratories

Hiroshi Ishiguro
Osaka University

Takayuki Kanda is a computer scientist with interests in intelligent robots and human-robot interaction; he is a researcher in the Intelligent Robotics and Communication Laboratories at ATR (Advanced Telecommunications Research Institute), Kyoto, Japan. **Takayuki Hirano** is a computer scientist with an interest in human–robot interaction; he is an intern researcher in the Intelligent Robotics and Communication Laboratories at ATR, Kyoto, Japan. **Daniel Eaton** is a computer scientist with an interest in human–robot interaction; he is an intern researcher in the Intelligent Robotics and Communication Laboratories at ATR, Kyoto, Japan. **Hiroshi Ishiguro** is a computer scientist with interests in computer vision and intelligent robots; he is Professor of Adaptive Machine Systems in the School of Engineering at Osaka University, Osaka, Japan, and a visiting group leader in the Intelligent Robotics and Communication Laboratories at ATR, Kyoto, Japan.

CONTENTS

1. INTRODUCTION
    1.1. Research on Partner Robots
    1.2. Social Relationships Over Time
    1.3. Technologies for Creating Human–Robot Relationships
2. FIELD TRIAL
3. METHOD
    3.1. Setting and Participants
    3.2. The Robot and Its Behavior
        Interactive Humanoid Robot “Robovie”
        Person Identification
        Interactive Behaviors
    3.3. Procedure
    3.4. Data Collection
        Time Spent Interacting With the Robot
        Tests of English Skills
        Social Interaction Around the Robot
4. RESULTS
    4.1. Preliminary Analyses
    4.2. Grade Differences
    4.3. Social Interaction
    4.4. Learning English
5. DISCUSSION
    5.1. Contributions to Human–Robot Interaction Methodology
    5.2. Contributions to the Theory of Human–Robot Interaction
    5.3. Contributions to the Design of Human–Robot Interaction
    5.4. Limitations
6. CONCLUSIONS

ABSTRACT

Robots increasingly have the potential to interact with people in daily life. It is believed that, based on this ability, they will play an essential role in human society in the not-so-distant future. This article examined the proposition that robots could form relationships with children and that children might learn from robots as they learn from other children. In this article, this idea is studied in an 18-day field trial held at a Japanese elementary school. Two English-speaking “Robovie” robots interacted with first- and sixth-grade pupils at the perimeter of their respective classrooms. Using wireless identification tags and sensors, these robots identified and interacted with children who came near them. The robots gestured and spoke English with the children, using a vocabulary of about 300 sentences for speaking and 50 words for recognition. The children were given a brief picture–word matching English test at the start of the trial, after 1 week and after 2 weeks. Interactions were counted using the tags, and video and audio were recorded. In the majority of cases, a child's friends were present during the interactions.

Interaction with the robot was frequent in the 1st week, and then it fell off sharply by the 2nd week. Nonetheless, some children continued to interact with the robot. Interaction time during the 2nd week predicted improvements in English skill at the posttest, controlling for pretest scores. Further analyses indicate that the robots may have been more successful in establishing common ground and influence when the children already had some initial proficiency or interest in English. These results suggest that interactive robots should be designed to have something in common with their users, providing a social as well as technical challenge.

1. INTRODUCTION

1.1. Research on Partner Robots

The development of humanoid robots such as Honda's ASIMO (Hirai, Hirose, Haikawa, & Takenaka, 1998) and interactive robots such as Sony's AIBO® (Fujita, 2001) and Kismet (Breazeal & Scassellati, 1999) has spawned a new area of research known as interactive robotics. These are not robots performing simple iterative tasks in factories or using specific tools in professional services such as surgical or military tasks (Thrun, 2004). Rather, this new wave of research is exploring the potential for *partner robots* to interact with people in daily life. Our research explores some fundamental problems in this new field.

Several researchers and companies have endeavored to realize robots as partners for people, and the concept of a partner robot is rapidly emerging. Typically equipped with an anthropomorphic body and various sensors used to interact with people naturally, the partner robot acts as a peer in everyday life. A humanoid robot, for example, guides office visitors by speech and with a hand-gesture recognition mechanism (Sakagami et al., 2002). For the home environment, NEC Corporation (2002) developed a prototype of a personal robot that recognizes individuals' faces, entertains family members with its limited speech ability, and performs as an interface to television and e-mail. Partner robots have also appeared in therapeutic applications. For example, Dautenhahn and Werry (2002) are applying robots to autism therapy. As these examples show, partner robots are beginning to participate in human society by performing a variety of tasks and functions.

Eliza was the first computer agent that established a relationship as a partner (Weizenbaum, 1966). People tried to interact with Eliza without necessarily having a specific task or request in mind. They sometimes made brief small talk and at other times engaged deeply in conversation. As Reeves and Nass (1996) discovered, humans unconsciously behave toward such a computer as if it were human. In recent robotics research, several pioneering studies have suggested that humans also can establish relationships with pet robots. Many people actively interact with animal-like pet robots. For example, people have adapted to the limited interactive ability of the robot dog, AIBO (Friedman, Kahn, & Hagman, 2003; Fujita, 2001). Furthermore, pet robots have been used successfully in therapy for the elderly, with some positive effects of their usage confirmed in long-term trials (Wada, Shibata, Saito, & Tanie, 2002).

1.2. Social Relationships Over Time

Recognizing the other person's identity, discovering similarities, and finding common ground are key issues in cementing social relationships. As Isaacs and Clark (1987) proposed, when people first meet, they gradually establish common ground through conversation. Empirical studies have shown that interlocutors adapt their speech to each other's attitudes and experience, weighing each other's perspectives when listening and making themselves understood (Fussell & Krauss, 1992). In forming satisfying and stable intimate relationships, they may even find similarities in their partner that do not exist in reality and tend to assume that their partner is a mirror of themselves (Murray, Holmes, Bellavia, Griffin, & Dolderman, 2002).

This evidence shows the importance of finding common ground in establishing relationships. However, relationships among people evolve over time (Hinde, 1988), and we believe people's attitude toward technological artifacts and their relationship with them also evolves over time. Little previous research has focused on long-term relations between individuals and computer systems in general or partner robots in particular. Short-term and long-term analyses must be carried out to evaluate partner robots. With respect to short-term experiments, many evaluation methods and systems have been proposed within the field of human–computer interaction and robotics. For instance, Quek et al. (2002) developed a gesture recognition-based system to analyze multimodal discourse. In robotics, Nakata, Sato, and Mori (1998) analyzed the effects of expressing emotions and intention. We have also performed several similar experiments, such as examining the effects of behavior pattern on impressions (Kanda, Ishiguro, & Ishida, 2001; Kanda, Ishiguro, Ono, Imai, & Nakatsu, 2002). However, in short-term human–robot interaction, we can only observe first impressions and the initial process of establishing relationships.

Some previous research has stressed the importance of long-term studies. Fish, Kraut, Root, and Rice (1992) evaluated a videoconferencing system and analyzed the transition of system use during 1 month of experimentation. Petersen, Madsen, and Kjær (2002) reported on the process of gaining experience with a new television system. These studies showed that the relation between human and agent is likely to change over time, just as interhuman relationships do. Therefore, it is vital to observe relationships between individuals and partner robots in an environment where long-term interaction is possible. The result of immersing a robot in an environment that demands ongoing participation is likely to be entirely different from that of exhibiting the robot in a public place like a museum, where the people who interact with it are transient.

1.3. Technologies for Creating Human–Robot Relationships

As previous research on interpersonal communication indicates, it is vital that two parties recognize each other for their relationship to develop. We cannot imagine having human partners or peers who cannot identify us. It is because we are able to identify individuals that we can develop a unique relationship with each of them (Cowley & MacDorman, 1995; Hinde, 1988). Although person identification (ID) is an essential requirement for a partner robot, current visual and auditory sensing technologies cannot reliably support it. Therefore, an unfortunate consequence is that a robot may behave the same with everyone.

Given only visual and auditory sensors, it is difficult to implement a person ID mechanism in robots that works in complex social settings. Many people may be talking at once, lighting conditions may vary, and the shapes and colors of the objects in the environment may be too complex for current computer vision technologies to function. In addition, the method of ID must be robust. Misidentification can ruin a relationship. For example, a person may be hurt or offended if the robot were to call the person by somebody else's name. To make matters worse, partner robots that work in a public place need to be able to distinguish between hundreds of people and to identify nearby individuals simultaneously. For instance, consider a situation involving people and robots working together in an office building, school, or hospital.

Besides their ability to identify and recognize others, robots should have sufficient interaction ability. In particular, human interaction largely depends on language communication. Whereas speaking is not so difficult for the partner robot, listening and recognizing human utterances is one of the most difficult challenges in human–robot interaction. Although some of the computer interfaces successfully employ speech input via microphone, it is far more difficult for the robots to recognize human utterances, because the robots suffer from noise from surrounding humans (background talk) and the robot body (motor noise). Little research has reported the solutions to this serious problem. We cannot expect ideal language perception ability like humans. However, we believe that robots can maintain interaction with humans, if they can recognize other human behaviors, such as distance, touching actions, and visual movements, in addition to utterances.

People have bodies that afford sophisticated means of expression through diverse channels. We believe that a robot partner, ideally, would have a humanlike body. A robot with a human-like body allows people to intuitively understand its gestures, which in turn causes people to behave unconsciously as if they were communicating with a human. These effects have even been observed with screen-bound agents that move and point (Isbister, Nakanishi, Ishida, & Nass, 2000). We believe that this anthropomorphic basis not only supports the embodiment of computer interfaces (Cassell et al., 1999), but also enables their grounding in social relationships (Cowley & MacDorman, 1995). Eye contact, gesture observation, and imitation in human–robot interactions greatly increase people's understanding of utterances (Ono & Imai, 2000). Close synchronization of embodied communication also plays an important role in establishing a communicative relation between the speaker and listeners (Ono, Ishiguro, & Imai, 2001). We believe that in designing an interactive robot, its body should be based on the human body to produce the most effective communication.

When partner robots are involved in people's daily life, they will take on certain roles and contribute to humans based on their skills. Apparently, a robot that is skilled at a single or limited set of tasks cannot satisfy the designation of partner. For example, a museum tour guide robot (Burgard et al., 1998) is equipped with robust navigational skills, which are crucial to its role; however, humans still do not perceive such a robot as their partner but see it merely as a museum orientation tool. What we recognize as a partner is probably a robot that can develop various kinds of relationships with humans. This does not mean simply performing multiple tasks. Rather, we believe that it is important to establish interactive relationships first, and then the tasks and skills of partner robots will gradually emerge along with advancing technologies.

2. FIELD TRIAL

Field trials provide an important means of exploring the potential of partner robots. We need extended observations because social relationships develop over time. In our field trial, two humanoid robots that had various communicative behaviors interacted with children at an elementary school. The purpose of the trial was for the robots to play with the children and to communicate with them in English, thus improving the children's ability to speak English. We observed the children's reactions to the robots over the course of 2 weeks. To the best of our knowledge, ours was the first extended trial using interactive humanoid robots in an authentic social setting.

Our choice of a task for the robot was motivated by the generally poor English language ability of Japanese people. We believe a lack of motivation and opportunities to speak English is a major cause of this deficiency. According to Gardner and Lambert (1972), the two main reasons for learning a second language are instrumental motivation (e.g., to earn a degree or obtain a position) and integrative motivation (e.g., to understand a different culture or to befriend foreigners). Many children in elementary and junior high school lack motivation and do not recognize the importance and usefulness of English. In fact, children have no need to speak English in Japan. Although English teachers speak English during class, children speak Japanese outside of class. In their daily lives, they almost never encounter foreigners who do not speak Japanese. Therefore, many children are not motivated to study English.

3. METHOD

We performed the field trial at an elementary school affiliated with Wakayama University. Two identical humanoid robots were put in the open corridor near the first- and sixth-grade classrooms, and for 2 weeks the two robots interacted with first-grade students and sixth-grade students. The following subsections describe the method of the trial in more detail.

3.1. Setting and Participants

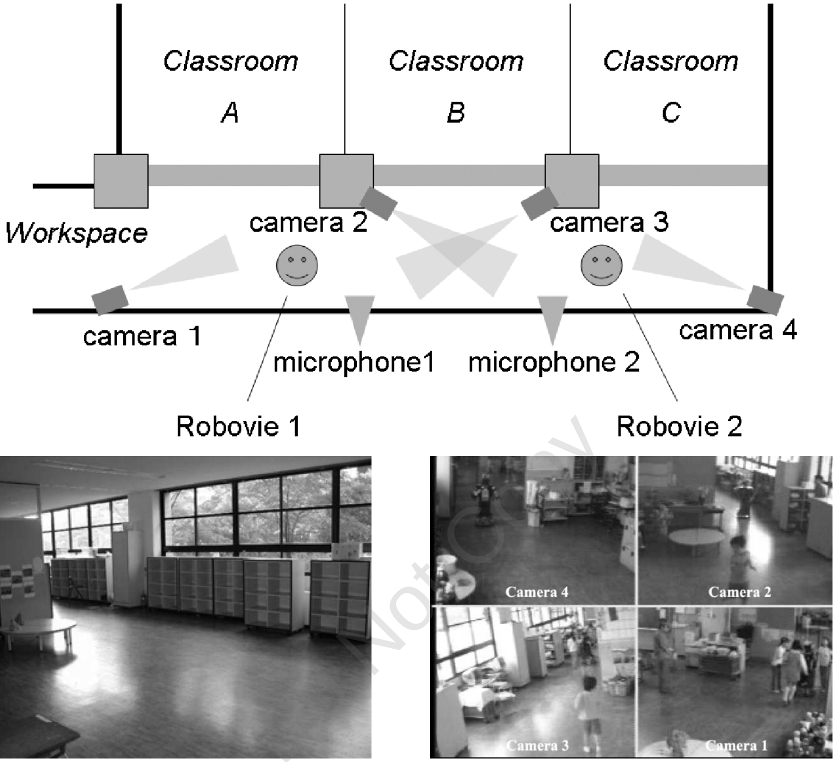
We carried out two sessions, one for first graders and the other for sixth graders. In general, there are six grades in a Japanese elementary school. This particular elementary school has three classes for each grade and about 40 students in each class. There were 119 first-grade students (6–7 years old; 59 boys and 60 girls) and 109 sixth-grade students (11–12 years old; 53 boys and 56 girls).

Figure 1 shows the three classrooms of the first grade. There are no walls between the classrooms and corridor, so that the corridor (called a workspace) is open to every first grader. The first graders' classrooms are located on the ground floor; the sixth graders' classrooms have the same layout as the first graders' and are located on the third floor.

3.2. The Robot and Its Behavior

The interactive humanoid robot we developed used a wireless ID tag system to identify different individuals. With visual, auditory, tactile, and wireless ID tag information, the robot took the initiative in interacting with children. For example, it called a child's name and initiated interaction after detecting the child from his or her ID tag. The robot could only recognize and speak English,

Figure 1. **Elementary school where robots were installed for 2 weeks, showing the open space plan of the school and where the cameras were installed.**

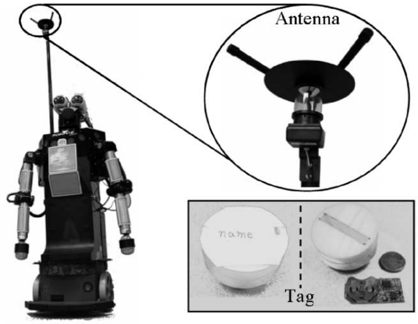


and its voice sounded somewhat like that of a child. The robot's utterances were based on recordings of a native English speaker (a professional narrator).

Interactive Humanoid Robot "Robovie"

Figure 2 shows the humanoid robot Robovie (Ishiguro, Ono, Imai, & Maeda, 2001). The robot is capable of human-like expression and recognizing individuals by using various actuators and sensors. Its body possesses highly articulated arms, eyes, and a head, which were designed to produce sufficient gestures to communicate effectively with humans. The sensory equipment includes auditory, tactile, ultrasonic, and vision sensors, which allow the robot to behave autonomously and to interact with humans. All processing and control systems, such as the computer and motor control hardware, are located inside the robot's body.

Figure 2. **Robovie and the wireless tag. Robovie (left) is an interactive humanoid robot that autonomously speaks, makes gestures, and moves around. With the antenna and tags, it is able to identify individuals.**

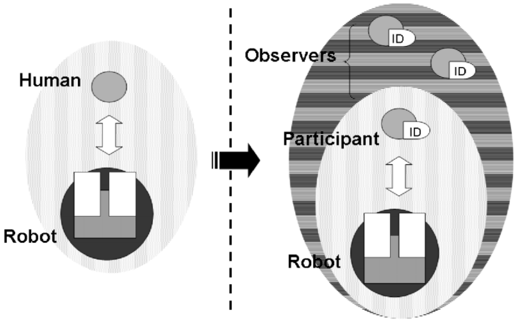


Person Identification

To identify individuals, we developed a multiperson ID system for partner robots by using a wireless tag system. Recent radio frequency ID (RFID) technologies enabled us to use contactless ID cards in practical situations. In this study, children were given easy-to-wear nameplates (5 cm in diameter) in which a wireless tag was embedded. A tag (shown in Figure 2, lower right) periodically transmitted its ID to the reader, which was onboard the robot. In turn, the reader relayed received IDs to the robot's software system. It was possible to adjust the reception range of the receiver's tag in real time from software. The wireless tag system provided the robots with a robust means of identifying many children simultaneously. Consequently, the robots could show some human-like adaptation by recalling the history of interaction with a given person.

The robot could also distinguish between participants and listeners. In linguistic research, Clark (1996) classified people in the process of communicating into two categories: participants and listeners. Participants speak and listen, whereas listeners are an audience. Based on Clark's theory, we modeled daily communication among children and a robot as shown in Figure 3. This model does not include distant communication, such as a member of an audience questioning a presenter at a speech. The left side of the figure shows a situation in which a robot could not identify individuals. The right side of the figure shows the people around the robot classified into two categories: participants and observers. The participant category is similar to Clark's definition of participant, but the observer category does not include eavesdroppers who listen in without the speaker's knowledge, because we are only concerned with the people within the robot's sensor range. Furthermore, we assume that the distance between the robot and people is adequate for the robot to distinguish between the two categories.

Figure 3. **The robot's communication model. The robot identifies multiple people simultaneously and classifies them into two categories for adapting its behaviors to them.**



Hall (1966) discussed several zones of proximity between humans in a conversation. According to his theory, a conversational distance is within 1.2 m, and a common social distance for people who have just met is between 1.2 m and 3.5 m. In this study, we defined the participant as the person who was within 1.2 m and nearest to the robot. This definition is based on our assumption that the participant in the communication process would approach the robot as they interact. In addition, a previous study showed the average distance in human–robot interaction was about 50 cm (Kanda, Ishiguro, Ono, Imai, & Nakatsu, 2002), which also supports the contention that the participant will keep within 1.2 m. Meanwhile, other individuals who stayed within the detectable range of the robot were considered to be observers, because they did not communicate with the robot but were within its region of awareness. A detailed mechanism and performance of the person ID system is described in Kanda, Hirano, Eaton, and Ishiguro (2003).

Interactive Behaviors

Robovie has a software mechanism for performing consistent interactive behaviors (Kanda, Ishiguro, Ono, Imai, & Mase, 2002). The intention behind the design of Robovie is that it should communicate at a young child's level. One hundred interactive behaviors have been developed. Seventy of them are interactive behaviors such as hugging (Figure 4), shaking hands, playing paper–scissors–rock, exercising, greeting, kissing, singing, briefly conversing, and pointing to an object in the surroundings. Twenty are idle behaviors such as scratching the head or folding the arms, and the remaining 10 are moving around behaviors. For the purpose of English education in this study, the situated module could only speak and recognize English. In total, the robot could utter more than 300 sentences and recognize about 50 words.

Figure 4. **Interactive behaviors of Robovie: (a) shake hands, (b) hug, (c) paper–scissors–rock, (d) exercise.**

Several interactive behaviors depended on the person ID function. For example, there was an interactive behavior in which the robot called a child's name if that child was at a certain distance. This behavior was useful for encouraging the child to come and interact with the robot. Another interactive behavior was a body part game; the robot asked a child to touch a part of the body by saying the part's name.

These interactive behaviors appeared in the following manner based on simple rules. The robot sometimes triggered the interaction with a child by saying, "Let's play, touch me," and it exhibited idling or moving-around behaviors until the child responded; once the child reacted, it continued performing friendly behaviors as long as the child responded to it. When the child stopped reacting, the robot stopped the friendly behaviors, said, "good bye," and restarted its idling or moving-around behaviors.

3.3. Procedure

Both sessions (for first and sixth grade) were conducted for 2 weeks, which is equivalent to 9 days of school. We gave the children safety instructions before the trial. Pictures of the robot were accompanied by messages in Japanese such as, "Do not treat the robots roughly," and "Do not touch the joints because it is not safe." We did not give the children any further instructions.

The two robots were put in the corridor as shown in Figure 1 (indicated as Robovie 1 and Robovie 2). The children were allowed to interact freely with both robots during recess. Every child had a nameplate with an embedded wireless ID tag (Figure 2, right bottom) so that the robots could identify the child during interactions.

The teachers were not involved in the field trial. Two experimenters (university students) looked after the two robots. They did not help the children interact with the robots but simply ensured the safety of the children and robots. For example, when the children crowded closely around the robot, the experimenters would tell them to maintain a safe distance.

3.4. Data Collection

Time Spent Interacting With the Robot

Each robot was equipped with a wireless ID tag reader that detected and identified ID tags embedded in the nameplates given to the children (described in Section 2). After identifying the children's IDs, the robot made a detection log of IDs for later analysis in addition to using it during interaction with the children. We prepared a simple program to calculate the interaction time per day for every child recorded in the detection log.

We also recorded scenes from the field trials with four cameras and two microphones. Figure 1 (upper and bottom right) describes the arrangement of the cameras and microphones and the obtained scenes of the trial. The video was used to verify the consistency of the wireless ID tag system. It was not analyzed otherwise.

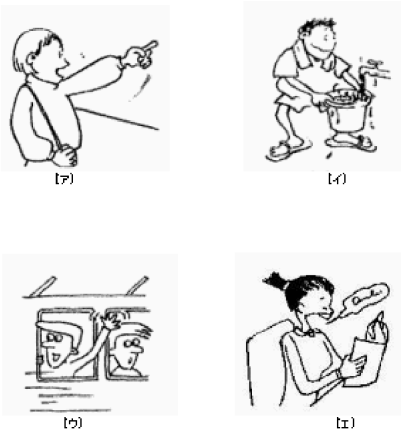
Tests of English Skills

The experimenters came to the first- and sixth-grade classes three times during the trial, and each time administered a brief English skills test: a pretest before the session, a test 1 week after the session began, and a posttest at the end of the 2-week session. Each test quizzed the students on the same six easy daily sentences used by the robots: "Hello," "Bye," "Shake hands please," "I love you," "Let's play together," and "This is the way I wash my face" (a phrase from a song), with the order of sentences changed in each test. We replayed the recorded voice of a native speaker for the test. On the answer sheets were four pictures for each phrase, and children had to choose the correct scene corresponding to the utterance (Figure 5). The score of the listening test for an individual was expressed as a percentage of the total number of correct answers, and thus the range of the listening test score always fell between 0 and 1.0.

Social Interaction Around the Robot

We also administered a questionnaire that asked the children to write down the names of their friends. These names were compared with the log data from

Figure 5. **Example of a question in the listening test: Which is "Bye"?**



the wireless tags to calculate the time children spent with the robot and their friends together.

4. RESULTS

We analyzed the effect of the robots on social interaction over time and learning by conducting quantitative statistical tests on the tag data and the English test scores.

4.1. Preliminary Analyses

In Figure 6, we describe the main measurements used in this study (i.e., the number of minutes each child interacted with the robot in the 1st and 2nd weeks of the trial; their English scores on the pretest, 1st week test, posttest; and the amount and percentage of time they interacted with the robot in the presence of friends). The figure contains the correlations among the main variables. There was no overall improvement in English scores among the students (although, as noted later, we found improvement among those who spent more time with the robot in the 2nd week).

4.2. Grade Differences

In Figure 7, we show how the children in the first and sixth grades interacted with their robot over the 2-week period. The figure shows that first graders spent more time interacting with the robot than sixth graders did, and the

Figure 6. **Descriptive statistics and correlations among the measures.**

Measures	*M*	*SD*	1	2	3	4	5	6
1. Child's interaction time with robot in minutes, 1st week (1)	12.5	14.0	1.00					
2. Child's interaction time with robot in minutes, 2nd week (2)	2.7	5.4	0.27	1.00				
3. Percentage interaction with friends (3)	67%	—	–0.11	–0.02	1.00			
4. Pretest English score	0.69	0.16	–0.02	–0.11	0.08	1.00		
5. English score after 1st week	0.70	0.16	–0.12	–0.13	0.03	0.37	1.00	
6. English score after 2nd week	0.69	0.16	–0.04	0.10	–0.05	0.35	0.40	1.00

Note. Correlations equal to or greater than ± .135 are significant at the .05 level or better.

Figure 7. **Interaction time with robots of first-grade students and sixth-grade students.**

	1st Week					2nd Week			
Grade	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1st[a]									
M (min.)	7.25	1.85	1.88	2.08	1.60	1.08	0.74	0.13	0.61
SD	7.36	3.57	3.14	4.90	3.77	3.00	2.43	0.51	2.35
6th[b]									
M (min.)	3.33	3.09	0.59	1.15	1.30	1.31	0.79	0.20	0.77
SD	5.15	5.94	2.01	2.87	2.74	2.64	2.48	0.88	1.37

[a] $n = 119$. [b] $n = 109$.

robot sustained their interest longer. It also indicates that the interaction between the children and the robots generally diminished in the 2nd week.

Nonetheless, a few children sustained a relationship with the robot. Child A said, "I feel pity for the robot because there are no other children playing with it," and Child B played with the robot for the same reason.

4.3. Social Interaction

We were surprised by the frequency with which children interacted with the robot in the company of other children (see Figures 8 and 9). Sixty-three percent of a first grader's interaction time with the robot was in the company of one or more friends. Seventy-two percent of a sixth grader's interaction time

Figure 8. **Scenes of the interactions between Robovie and students. (a) First-grade students with the robot on Day 1. (b) First-grade students with the robot during the 2nd week. (c) Sixth-grade students on Day 1. (d) Sixth-grade students during the 2nd week.**

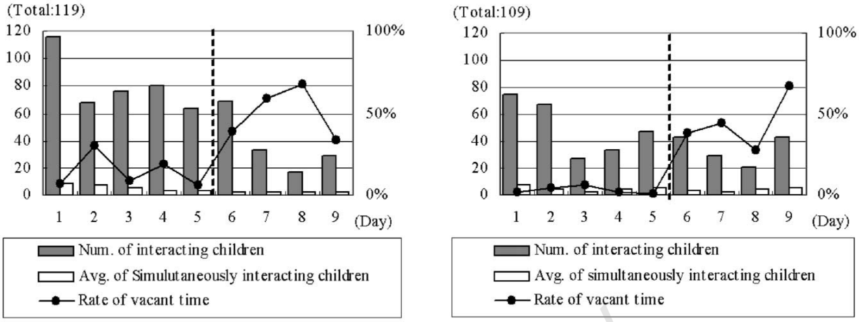
(a)

(b)

(c)

(d)

***Figure 9.* Transition in number of children playing with the robot. (a) Results for first-grade students. (b) Results for sixth-grade students. Number of interacting children represents the total number of the children identified by each robot's wireless system each day. Average of simultaneously interacting children represents the average number of children who simultaneously interacted with the robot. Rate of vacant time is the percentage of the time there was no child around the robot during each day.**



(a) (b)

with the robot was in the company of one or more friends. Because the presence of friends could have affected each child's learning of English from the robot, we controlled for the presence of friends in the following analyses.

4.4. Learning English

The analyses we present are analyses of variance in which the dependent variable is the improvement in each child's English test score from the child's English pretest score. Although many children did not know English at the beginning of the trial, some knew a bit. If they knew any of the phrases on the English test (such as "bye") their improvement might have been small owing to a ceiling effect. Therefore, the appropriate analysis of the effects of the robot on learning is the change from the pretest to the posttest, controlling for the initial pretest score. The main analyses we ran were standard least squares analyses, described as follows:

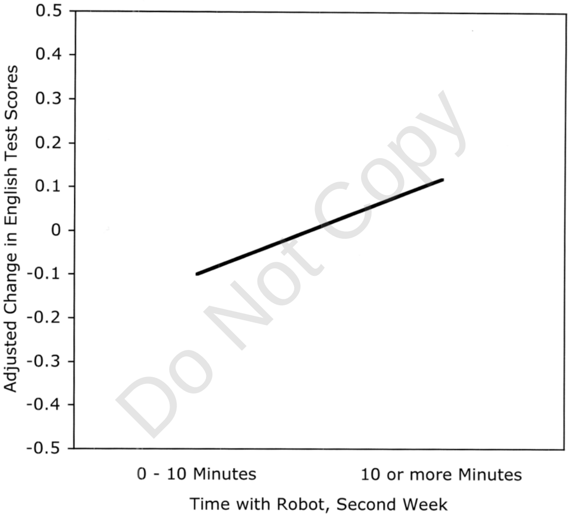
Model (2nd week English score – pretest English score) = intercept + pretest English score + Week 1 interaction minutes with robot + Week 2 interaction minutes with robot + percentage of interaction time with friends.

The results of this analysis are shown in Figures 10 and 11. This analysis showed the expected significant ceiling effect of pretest English scores on the change in scores from pretest to posttest, $F(1, 198) = 86$, $p < .001$. That is, the more English the children already knew at the beginning of the trial, the less they learned from the robot. However, the amount of time they interacted

Figure 10. **Analysis of variance results for effect of interaction time with the robot on improvement in English scores at the posttest (after 2 weeks).**

Source	**df**	***F* Ratio**	***p***
Pretest English test score	1	85.8	$<.0001$
Percentage of interaction time with friends	1	1.5	*ns*
Interaction time with robot, 1st week	1	1.4	*ns*
Interaction time with robot, 2nd week	1	5.6	.019
Error	198		

Figure 11. **Change in English score as a function of interaction time with the robot in the 2nd week, controlling for pretest English score and presence of friends.**



with friends and the robots together did not have an impact on the change in the English scores. The amount of time children spent with the robot during the 1st week also had no effect on their improvement in English by the 2nd week, but the amount of time that children interacted with the robots during the 2nd week did have a significant and positive impact on improvement in English in the 2nd week, $F(1, 198) = 5.6$, $p = .02$, $d = .33$.

Because we found significant improvement in English learning after 2 weeks, we examined whether there was any evidence of improvement after only 1 week with the robot. This analysis showed that time spent with the robot during the 1st week did not have a significant impact on the change in the English scores from the pretest to the 1st week's scores. Indeed, the trend was slightly negative ($p < .10$). The absence of a 1st-week result suggests that learning depended on a sustained interest in the robot and maintaining a relationship with it. It was only those children who continued to interact with the robot through the 2nd week—those who formed a relationship with the robot—who learned from it.

We also investigated whether the grade of the children influenced the improvement in the English scores. To examine this we added the grade (first or sixth) to the previous equation, and included statistical interactions of grade with the presence of friends, time with the robot in the 1st week, and the presence of friends in the 2nd week. The results of this analysis did not change the overall positive effect of interaction time with the robot in the 2nd week (i.e., the relationship shown in Figure 10). However, this analysis did show that sixth graders learned more English than first graders ($p < .01$), and that first graders benefited slightly more from interaction with the robot in the 1st week ($p < .08$).

One alternative explanation to the improvement in English scores at the end of 2 weeks is that causality was reversed. That is, perhaps those children who were more interested in English and knew more English at the start of the trial were more interested in interacting with the robot. To investigate the possibility that knowledge of English caused the children to interact with the robot more, we ran a regression analysis examining the impact of pretest English scores on 1st and 2nd week of time spent with the robot, controlling for the presence of friends. Pretest English scores did not predict the 1st week of time with the robot, but there was a marginal positive effect of pretest scores on time with the robot in the 2nd week, $F(1, 208) = 2.5$, $p = .11$. This suggests that part of the reason for the results shown in Figures 10 and 11 might be the initial ability of some children to understand the robot's English and feel comfortable with it. They might have felt they had something in common with the robot (i.e., the English language).

5. DISCUSSION

We believe that this field trial provided us with many useful insights that we can apply to the development of future partner robots. The humanoid robots autonomously interacted with children by using their human-like bodies and various sensors such as visual, auditory, and tactile sensors. They also had a mechanism to identify individuals and to adapt their interactive behaviors to them.

The results suggest that the robot did encourage some children to improve their English, and that the robot was more successful in engaging children who already knew at least a little English. These findings support arguments based on previous literature in social psychology on similarity and common ground;

they suggest robots should be designed to have attributes and knowledge in common with their users.

5.1. Contributions to Human–Robot Interaction Methodology

Our results suggest that the impact of the robot did not show up until the 2nd week. This finding supports the argument that a robot's influence will depend on its ability to create a relationship with the user. It also suggests that a robot's effect on individuals changes over time. Therefore, we need to study long-term interactions to learn how to create effective partner robots.

5.2. Contributions to the Theory of Human–Robot Interaction

Our field trial highlighted the important unsolved aspects of human–robot interaction in an authentic social setting. The trial showed gradual loss of interest in interacting with the robot among most of the children. It was an important finding that the children interacted with the robot for the duration of 1 week; however, our robots failed to keep most of the children's interest after the 1st week. We believe that the robots' first impact created unreasonably high expectations in the children. The children mobbed the robot, overwhelming its ability to interact. In other words, the robot could not cope with the children's enthusiasm. Although partner robots are making news in Japan (such as Honda's and Sony's humanoid robots, and the big exhibition on partner robots named ROBODEX; ROBODEX Executive Committee, 2003), the robots' ability to be a partner to people is still lacking. Robots are very novel in general; therefore, their first impact can induce a greater desire for communication than their interactive ability can satisfy. In our trial, the children's interaction with the robots gradually decreased, especially during the 2nd week. Therefore, our trial showed us the limitation of the robots' ability to maintain long-term relationships and the disappointment that followed the robot's initial impact. However, we believe unreasonable expectations will diminish as partner robots become commonplace.

Regarding the body and appearance of the robot, our results seem to encourage the use of a humanoid robot. We believe that the body of a humanoid robot played a useful part in establishing common ground. That is, a robot that possesses a humanoid body will be more successful at sustaining interaction because people see it as similar to themselves and that it interacts as they do. Nonetheless, we need further research to establish a model of these kinds of social effects, such as common ground to see if they are more easily achieved with a humanoid robot by comparing humanoid and nonhumanoid robots.

It is also necessary to formalize a model of the relationships between humans and robots over time, and establish a method to promote lasting interactive relationships. Several pet robots have a special pseudolearning mechanism: Although they have many functions, they only show a few functions at first and then gradually reveal more according to their interactions. Furthermore, if robots really learn something about an individual to personalize the relationship, the robots will be able to build closer relationships with people. Therefore, identifying and defining the mechanism for sustaining long-term relationships is an important area of future research in human–robot interaction.

5.3. Contributions to the Design of Human–Robot Interaction

The trial showed that with respect to the interactive ability based on sensor data processing, real-world data are vastly different from that produced in a well-controlled laboratory. For example, many children ran around and spoke loudly to the robot; thus, its speech recognition was not effective in the classroom where the trial was carried out. To design robots that operate in real-world settings, we must consider how to make sensing more robust. Although many researchers and developers have been developing and improving sensing technologies, such as vision processing and speech recognition, robots still have weak ability compared to that of humans.

Fortunately, the wireless ID of persons worked well in our trial. We believe that one of the potentially promising approaches for acquiring interactive ability in the real world is to use environment-based sensors such as the wireless ID tags. In the trial, we observed several positive effects of the ID tags on the children's interaction with the robots:

- Child C did not seem to understand English at all. However, once she heard her name uttered by the robot, she became quite pleased and began interacting more frequently with the robot.
- Children D and E counted how many times the robot called their respective names. D's name was called more often, so D proudly told E that the robot preferred D.
- Child F passed by the robot. He did not intend to play with the robot, but because he saw Child G playing with the robot, he joined in.

These examples suggest that person ID was one of the triggers of the interaction and an essential behavior for continuous interaction.

Our robot currently recognizes only those who are around it. That is, even if the robot is faced with multiple parties, it does not distinguish the relationships among them. However, as the previous example indicates, relationships among people might affect the interaction. For example, a child may take a friend to the robot, or someone may take part in the interaction because a friend is playing with the robot. Therefore, we believe a partner robot should also recognize relationships between children (friendship, hostility, etc.).

5.4. Limitations

This study was a field trial rather than a true experiment with controls. For example, we did not compare the robot with an ordinary computer English teaching game. A detailed experiment might offer more precise and reliable results on the teaching of English to Japanese students. However, our main goal was not to teach English optimally but to learn how to create partnership in a robot. We believe that field trials in a frontier research area (e.g., partner robotics) are essential for developing the discipline. A field trial provides us with valuable information on the deficiencies in our approach, which is helpful to inspire future technological developments. We would be pleased if this work inspired rigorous research in the social aspects of human–robot interaction.

We did not associate videotaped interactions with tag data from each child. We believe this kind of fine-grained analysis would be particularly useful, for example, in checking the number of utterances of each child and in observing how each child initiated interaction with the robot. In future research, it would be very helpful to code the videotapes and thereby combine qualitative observations with tag data, which lack detailed information.

6. CONCLUSIONS

We performed a field trial for 2 weeks using interactive humanoid robots with first- and sixth-grade elementary school students. In the trial, the robots behaved as English peer tutors for Japanese students. The results suggest that the robot did encourage some children to improve their English. Our findings demonstrate the possibility of having interactive robots work in our daily life, although the benefits may be still too small to justify practical application. If the interactive robots were to acquire a more powerful ability to maintain relationships with humans, we would feel more confident in them serving various roles in our daily life in the immediate future. This result would encourage further robotics and human–computer interaction research related to sociality (e.g., theory of common ground), expression ability including humanoid control, sensory and recognition ability, and more metalevel communication mechanisms.

NOTES

Acknowledgments. We thank the teachers and students at the elementary school for their eager participation and helpful suggestions. We also thank the reviewers for their advice and help on this article.

Support. This research was supported by TAO Japan.

Authors' Present Addresses. Takayuki Kanda, Intelligent Robotics and Communication Laboratories, ATR, 2–2-2 Hikaridai, Soraku-gun, Seika-cho, Kyoto, 619-0224, Japan. E-mail: kanda@atr.co.jp. Takayuki Hirano, Intelligent Robotics and Communication Laboratories, ATR, 2–2-2 Hikaridai, Soraku-gun, Seika-cho, Kyoto, 619-0224, Japan. E-mail: t-hirano@atr.co.jp. Daniel Eaton, Intelligent Robotics and Communication Laboratories, ATR, 2–2-2 Hikaridai, Soraku-gun, Seika-cho, Kyoto, 619-0224, Japan. E-mail: deaton@atr.co.jp. Hiroshi Ishiguro, Adaptive Machine Systems, Graduate school of Engineering, Osaka University, 2–1 Yamadagaoka, Suita, Osaka, Japan. E-mail: ishiguro@ams.eng.osaka-u.ac.jp.

HCI Editorial Record. First manuscript submitted to ACM *Transactions on Computer Human Interaction* on February 3, 2003. Transferred to *Human-Computer Interaction* on June 17, 2003. Revision received July 28, 2003. Accepted by Sara Kiesler and Pamela Hinds. Final manuscript received September 1, 2003. — *Editor*

REFERENCES

Breazeal, C., & Scassellati, B. (1999). A context-dependent attention system for a social robot. *Proceedings of International Joint Conference on Artificial Intelligence.* San Francisco, CA: Kaufmann.

Burgard, W., Cremers, A. B., Fox, D., Hähnel, D., Lakemeyer, G., Schulz, D., et al. (1998). The interactive museum tour-guide robot. *Proceedings of 15th National Conference on Artificial Intelligence.* Menlo Park, CA, AAAI.

Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjalmsson, H., et al. (1999). Embodiment in conversational interfaces: Rea. *Proceedings of the CHI 99 Conference on Human Factors in Computing Systems.* New York: ACM.

Clark, H. H. (1996). *Using language.* Cambridge, England: Cambridge University Press.

Cowley, S. J., & MacDorman, K. F. (1995). Simulating conversations: The communion game. *AI & Society, 9,* 116–137.

Dautenhahn, K., & Werry, I. (2002). A quantitative technique for analyzing robot–human interactions. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems.* New York: IEEE.

Fish, R. S., Kraut, R. E., Root, R. W., & Rice, R. E. (1992). Evaluating video as a technology for informal communication. *Proceedings of the CHI 92 Conference on Human Factors in Computing Systems.* New York: ACM.

Friedman, B., Kahn, P. H., Jr., & Hagman, J. (2003). Hardware companions?—What online AIBO discussion forums reveal about the human–robotic relationship.

Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems. New York: ACM.

Fujita, M. (2001). AIBO: Towards the era of digital creatures. *International Journal of Robotics Research, 20,* 781–794.

Fussell, S., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality & Social Psychology, 62,* 378–391.

Gardner, R. C., & Lambert, W. E. (1972). *Attitude and motivation in second language learning*. Rowley, MA: Newbury House.

Hall, E. (1966). *The hidden dimension*. New York: Anchor.

Hinde, R. A. (1988). *Individuals, relationships and culture: Links between ethology and the social sciences*. Cambridge, England: Cambridge University Press.

Hirai, K., Hirose, M., Haikawa, Y., & Takenaka, T. (1998). The development of the Honda humanoid robot. *Proceedings of the IEEE International Conference on Robotics and Automation*. New York: IEEE.

Isaacs, E., & Clark, H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General, 116,* 26–37.

Isbister, K., Nakanishi, H., Ishida, T., & Nass, C. (2000). Helper agent: Designing an assistant for human–human interaction in a virtual meeting space. *Proceedings of CHI 2000 Conference on Human Factors in Computing Systems*. New York: ACM.

Ishiguro, H., Ono, T., Imai, M., Maeda, T., Nakatsu, R., & Kanda, T. (2001). Robovie: An interactive humanoid robot. *International Journal of Industrial Robots, 28,* 498–503.

Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2003). Person identification and interaction of social robots by using wireless tags. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. New York: IEEE.

Kanda, T., Ishiguro, H., & Ishida, T. (2001). Psychological analysis on human–robot interaction. *Proceedings of the IEEE International Conference on Robotics and Automation*. New York: IEEE.

Kanda, T., Ishiguro, H., Ono, T., Imai, M., & Mase, K. (2002). A constructive approach for developing interactive humanoid robots. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. New York: IEEE.

Kanda, T., Ishiguro, H., Ono, T., Imai, M., & Nakatsu, R. (2002). Development and evaluation of an interactive humanoid robot "Robovie." *Proceedings of the IEEE International Conference on Robotics and Automation*. New York: IEEE.

Murray, S. L., Holmes, J. G., Bellavia, G., Griffin, D. W., & Dolderman, D. (2002). Kindred spirits? The benefits of egocentrism in close relationships. *Journal of Personality & Social Psychology, 82,* 563–581.

Nakata, T., Sato T., & Mori, T. (1998). Expression of emotion and intention by robot body movement. *Proceedings of the IAS-5 International Conference on Intelligent Autonomous Systems*. Amsterdam, The Netherlands: IOS.

NEC Corporation. (2002). *Personal robot PaPeRo*. Retrieved September 1, 2003, from http://www.incx.nec.co.jp/robot/PaPeRo/english/p_index.html

Ono, T., & Imai, M. (2000). Reading a robot's mind: A model of utterance understanding based on the theory of mind mechanism. *Proceedings of 17th National Conference on Artificial Intelligence*. Menlo Park, CA: AAAI.

Ono, T., Ishiguro, H., & Imai, M. (2001). A model of embodied communications with gestures between humans and robots. *Proceedings of 23rd Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Petersen, M. G., Madsen, K. H., & Kjær, A. (2002). The usability of everyday technology: Emerging and fading opportunities. *ACM Transactions on Computer–Human Interaction, 9*, 74–105.

Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X. F., Kirbas, C., et al. (2002). Multimodal human discourse: Gesture and speech. *ACM Transactions on Computer–Human Interaction, 9,* 171–193.

Reeves, B., & Nass, C. (1996). *The media equation*. Cambridge, England: Cambridge University Press.

ROBODEX Executive Committee. (2003). *ROBODEX2003 (robot dream exposition)*. Retrieved September 1, 2003, from http://www.robodex.org/

Sakagami, Y., Watanabe, R., Aoyama, C., Matsunaga, S., Higaki, N., & Fujimura, K. (2002). The intelligent ASIMO: System overview and integration. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. New York: IEEE.

Thrun, S. (2004). Toward a framework for human–robot interaction. *Human–Computer Interaction, 19,* 9–24. [this special issue]

Wada, K., Shibata, T., Saito, T., & Tanie, K. (2002). Analysis of factors that bring mental effects to elderly people in robot assisted activity. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. New York: IEEE.

Weizenbaum, J. (1966). Eliza: A computer program for the study of natural language communication between man and machine. *Communications of the ACM, 9*(1), 36–45.