Covert Robot-Robot Communication: Human Perceptions and Implications for Human-Robot Interaction

Tom Williams, Priscilla Briggs, and Matthias Scheutz
Human-Robot Interaction Laboratory, Tufts University

As future human-robot teams are envisioned for a variety of application domains, researchers have begun to investigate how humans and robots can communicate effectively and naturally in the context of human-robot team tasks. While a growing body of work is focused on human-robot communication and human perceptions thereof, there is currently little work on human perceptions of robot-robot communication. Understanding how robots should communicate information to each other in the presence of human teammates is an important open question for human-robot teaming. In this paper, we present two human-robot interaction (HRI) experiments investigating the human perception of *verbal* and *silent* robot-robot communication as part of a human-robot team task. The results suggest that silent communication of task-dependent, human-understandable information among robots is perceived as creepy by cooperative, co-located human teammates. Hence, we propose that, absent specific evidence to the contrary, robots in cooperative human-robot team settings need to be sensitive to human expectations about overt communication, and we encourage future work to investigate possible ways to modulate such expectations.

Keywords: Joint human-robot teams, mixed initiative, robot-robot communication, uncanny actions, human perceptions of robot communication

1. Introduction

An important goal of human-robot interaction (HRI) is to develop methods for effective, natural interactions between humans and robots. While much research in HRI toward this goal has focused on the effects of robot appearance and observable behavior, a significant aspect of *natural HRI* is communication in natural language (e.g., Scheutz, Schermerhorn, Kramer, & Anderson, 2007), which has only recently received significant attention. Recent research has investigated various social aspects of natural language interactions with robots, such as politeness (e.g., Briggs & Scheutz, 2014), turn taking (e.g., Nadel, Revel, Andry, & Gaussier, 2004), affective speech (e.g., Scheutz, Schermerhorn, & Kramer, 2006), dialogue-appropriate facial movements (e.g., Liu, Ishi, Ishiguro, & Hagita, 2012), pragmatic analysis (e.g., Williams, Briggs, Oosterveld, & Scheutz, 2015), and collaborative control (e.g., Fong, Thorpe, & Baur, 2003). Due to the difficulty of managing multi-party dialogue, this research

Journal of Human-Robot Interaction, Vol. 4, No. 2, 2015, Pages 24-49. DOI 10.5898/JHRI.4.2.Williams

has primarily focused on dialogue between a single human and a single robot, with a few exceptions: For example, some work has demonstrated multiple simultaneous conversations between a single human and several remote robots (Fong et al., 2003), and some work has demonstrated multi-party conversation between multiple co-located humans and a single robot (Foster et al., 2012; Matsuyama, Taniyama, Fujie, & Kobayashi, 2006).

However, little research to date has investigated the question of how robots that communicate with humans should communicate with *each other*. Some researchers have looked into wireless inter-robot communication protocols (e.g., Balch & Arkin, 1994; Fukuda & Sekiyama, 1994; Wang, 1994), and some researchers have developed mechanisms for managing conversation between a human and multiple co-located robots (e.g., Briggs & Scheutz, 2012), but such research does not examine how humans actually perceive such communication. *Should* robots communicate with each other in natural language, so as to be transparent to humans, or can they use whatever form of communication best suits their needs?[1] It seems clear that there cannot be a simple context-independent answer to this question. For example, consider the difference between cooperative vs. competitive contexts: In the first, humans and robots have to work together toward a common set of goals; in the second, humans and robots have competing, incompatible goals. Socially assistive robots and robots for search and rescue missions are examples of the former, while robots for robo-soccer or law enforcement are examples of the latter. It is clear that in the latter case, robots should not divulge their intentions and goals, as leaking knowledge about their plans and actions will benefit the adversary. It is less clear whether in the first case robots should *always* communicate in natural language. In some instances, keeping co-present humans "in the loop" will be advantageous, while in others, "communications overhead" might be unnecessary and distracting.

In this paper, we set out to investigate human perceptions of robot-robot communication in the context of a mixed-initiative human-robot team, where the human commands two robots to perform a search and rescue task in a simulated disaster area. The main results of a set of two human-robot experiments in this domain suggest that robots might have to communicate in natural language with humans in the context of cooperative tasks in order to avoid being viewed as unsettling or creepy by their cooperative, co-located human teammates. The rest of the paper proceeds as follows: In Section 2, we briefly review previous work on human-robot and robot-robot communication, and then our hypotheses are described in Section 3. Sections 4 and 5 present the two HRI experiments we conducted to evaluate our four hypotheses. Section 6 discusses the implications of our findings, followed by a summary in Section 7.

2. Background

While there is a growing body of research on human perceptions of human-robot communication, very little work has investigated human perception of robot-robot communication. Two sets of studies have investigated verbal robot-robot communication by examining human perceptions of robots engaged in humorous banter or non-task-oriented conversation (Hayashi, Kanda, Miyashita, Ishiguro, & Hagita, 2008; Tsujimoto, Munekat, & Ono, 2013), and two sets of studies have investigated human perception of nonverbal robot-robot communication. As we are primarily concerned with human perception of silent robot-robot communication relative to verbal robot-robot communication, we will focus on these two studies.

[1] The authors find robots using the most effective means of communication to be non-controversial when humans are not present, telepresent, or otherwise in observation.

In the first set of studies (Kanda, Ishiguro, Ono, Imai, & Mase, 2002; Kanda, Ishiguro, Ono, Imai, & Nakatsu, 2004), human participants observed two robots discussing a piece of artwork. The robots' manner of conversation fit one of three conditions: In the first condition, the robots conversed and gesticulated; in the second, they conversed without gesticulating; and in the third, the conversation was skipped altogether. In all three conditions, one of the two robots subsequently approached and spoke to the human observer. The human subjects were then asked about their comfort level when interacting with the robot. No adverse effects were found, suggesting that it is perfectly acceptable for robots to converse silently while observed by humans.

However, there are two important limitations to this study. For one, the experiment does not truly contrast verbal and silent behavior, as in the silent condition, no robot-robot conversation whatsoever took place from the participant's point of view. It would thus be more accurate to say that the study compares the comfort levels of participants who engage in conversation with robots that have been shown capable of conversation, and the comfort levels of participants who engage in conversation with robots that have *not* been shown capable of conversation. Moreover, participants in the experiment had no investment in the robots' conversation; the robots were not discussing anything the participants needed to know about, and thus there were no negative consequences to participants being kept "out of the loop" of the robots' conversation. In a human-robot team task, information communicated between robots could very well be crucial for human teammates.

In the second set of studies investigating human perception of nonverbal robot-robot communication (Fraune & Šabanović, 2014a, 2014b), participants completed surveys while robot activities in their vicinity unfolded according to one of four conditions: (1) three robots wandered pseudo-randomly, beeping occasionally; participants were told that the robots did not communicate with each other, (2) three robots wandered pseudo-randomly, beeping occasionally; participants were told that the robots communicated with each other over the Internet, (3) three robots wandered pseudo-randomly, beeping occasionally, and from time to time beeping loudly in sequence; participants were told that the robots communicated via beeps, and (4) a control condition with no robots present. The researchers were interested in whether the attribution of non-anthropomorphic communication styles to the robots would increase the salience of the robots' "out-group status," causing them to be viewed less favorably. Results showed that participants generally thought the robots were communicating aloud, even in conditions 1 and 2, where participants where either told that the robots were not communicating or that the robots were communicating over the Internet. Since no significant differences in human perception of robots among any of the four conditions were found, the researchers concluded that the robots were not attributed out-group status, and that communication style did not affect human perceptions of robots.

However, as with the previous set of studies, there are two important aspects of this study which significantly limit its applicability to other robot-robot communication scenarios, in particular, human-robot team tasks. First, it is not clear whether there was any reason for the participants to have felt left "out of the loop," or to have felt that the robots were uncooperative, untrustworthy, or unsettling, as the participants did not know what the robots were doing and the robots never communicated verbally. Had the subjects been given the opportunity to observe the robots communicate verbally, then the use of silent communication could have been cast as an *intentional choice* of the robots to prevent the humans from knowing the content of their communications. Furthermore, the lack of verbal communication may have reduced the degree to which humans perceived the robots as human-like, thus decreasing the effects communication style may have had on perception of intention-driven robot attributes, such as cooperativity. As with the first set of studies, the human observers had no investment in the robots' activities, with all the consequences previously described.

The above sets of studies are typical of a whole class of experiments in HRI, where humans are interaction observers rather than participants, and as such, they have no reason to be invested in the robots' activities or performance. Hence, any conclusions derived from such experiments are limited to interaction observation and cannot automatically be generalized to interaction participation.

3. Human Perceptions of Covert Robot Communication

Human subjects will have far lower investment in communication outcomes between robots they are merely *observing* compared to robots with which they are *interacting*. To address this lack of investment, we devised a joint human-robot team task where human participants (1) have to interact verbally with robots, (2) are able to verify when silent communication has occurred, and (3) have a vested interest in the accuracy of the robot-robot communication. To ensure that the participants would have an interest in the information communicated, we constructed a scenario in which participants needed one robot to relay instructions to another robot. In this way, participants depended on the robot interlocutor to communicate their instructions accurately to the other robot (in order for the scenario's task to be completed efficiently), and the robots depended on the participant to provide them with appropriate instructions. This paradigm allowed us to explore four important questions about robot-robot communication in human-robot team tasks:

1. Will robots be viewed as more or less **trustworthy** if they choose to communicate silently? A wide variety of factors can influence the degree of trust a human has for a robot teammate. One such factor is transparency: To engender trust, the motivation behind a robot's behavior should be transparent and easily understandable (Hancock, Billings, & Schaefer, 2011). If robot-robot communication is enacted silently, the motivation behind robot actions may be unclear, leading to distrust. Another factor influencing human-robot trust is similarity of mental models; to engender trust, teammates should endeavor to create and share mental models (Hancock et al., 2011; Neerincx, 2007). If robot-robot communication is enacted silently, human teammates may not be able to appropriately update their mental models. The resulting dissimilarity of mental models may lead to distrust. Given these concerns, we hypothesize that **(H1)** robots will be viewed as **less trustworthy** if they choose to communicate silently.

2. Will robots be viewed as more or less **cooperative** if they choose to communicate silently? We believe that the same factors that may cause a robot to be viewed as untrustworthy may also cause a robot to be viewed as uncooperative, as lack of transparency and dissimilarity in mental models are likely to lead to simple misunderstandings. For this reason, we hypothesize that **(H2)** robots will be viewed as **less cooperative** if they choose to communicate silently.

3. Will robots be viewed as more or less **unsettling** if they choose to communicate silently? Over the past few decades, a variety of fields have given increased attention to the "Uncanny Valley" (Mori, MacDorman (Translator), & Minato (Translator), 2005), a hypothesis stating that entities very close to (but not quite) human are perceived as creepy or unsettling. Recent research suggests that these feelings of eeriness do not directly correlate with human-likeness, and that human likeness may thus be only one of several factors contributing to the Uncanny Valley effect (Brenton, Gillies, Ballin, & Chatting, 2005; MacDorman, 2006). One such contributing factor is the use of "uncanny actions." Uncanny

actions include those that can be construed as human but are executed with slight deviation from normal human execution: a robot that blinks too infrequently or that follows teammates too closely could be viewed as uncanny. In addition to these types of uncanny actions, we believe that actions that cannot be construed as human should also be considered to be uncanny actions, as research has shown that humans generally prefer robots whose actions can be construed as human (Walters, Syrdal, Dautenhahn, Te Boekhorst, & Koay, 2008). One example of this kind of uncanny action is *telepathy*. Telepathy is not in the realm of human ability and is largely considered to be paranormal or supernatural. However, robots regularly communicate in a manner reminiscent of telepathy (i.e., using wireless communication). This behavior may thus be perceived as creepy or unsettling. It is possible that this behavior would be viewed as analogous to text-messaging or other electronic forms of communication, or it could be viewed as analogous to situations in which humans seem to "guess" what their interlocutor is going to say (e.g., when couples finish each other's sentences). However, in such situations, there is an assumption that information could be "guessed" due to contextual factors or longitudinal learning of an agent's goals and preferences; whereas in silent robot-robot communication, information may be communicated for which it would be next to impossible for a robot to "guess" that information. We thus hypothesize that **(H3)** robots will be viewed as **more unsettling** if they choose to communicate silently.

4. Will robots be viewed as more or less **efficient** if they choose to communicate silently? Research suggests that the use of nonverbal cues in human-robot communication leads to higher efficiency (Breazeal, Kidd, Thomaz, Hoffman, & Berlin, 2005). We believe that humans will be able to recognize the efficiency inherent in completely non-verbal robot-robot communication. For this reason, we hypothesize that **(H4)** robots will be viewed as **more efficient** if they choose to communicate silently.

We will now introduce the details of the experimental paradigm first described in (Williams, Briggs, Pelz, & Scheutz, 2014), which we used to investigate these four questions.

4. Experiment 1

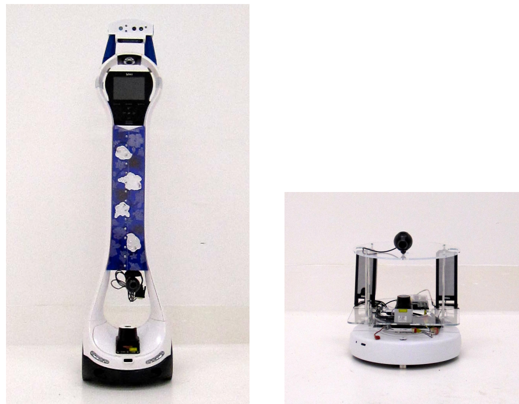
We employed a team task in which a human commander had to verbally assign different tasks to two robots and observe the robots' execution of those tasks, in order to accomplish the task goals.

4.1 Equipment

We used two different robots: "VGo" (Fig. 1a), a VGo telepresence robot augmented with an on-board computer and a variety of sensors (Tsui et al., 2013), and "Roompi" (Fig. 1b), an iRobot Create augmented with a Raspberry Pi computer, Hokuyo Laser Range Finder, speakers, and webcam. As the VGo is limited to a single text-to-speech voice option, we used that voice for both robots. While this voice was, in our opinion, slightly more female-sounding than male-sounding, it was the only option available. Both robots were controlled through Wizard of Oz interfaces, teleoperated by trained confederates in a nearby room.

4.2 Procedure

Participants were told that their task was to instruct robots as part of training for a disaster relief scenario, and that the adjacent room, which was filled with a number of boxes and other obstacles, simulated a power plant strewn with debris after a nuclear disaster. Participants were told that the sensors of the robots they would be instructing had been manipulated such



(a) “VGo,” the VGo telepresence robot

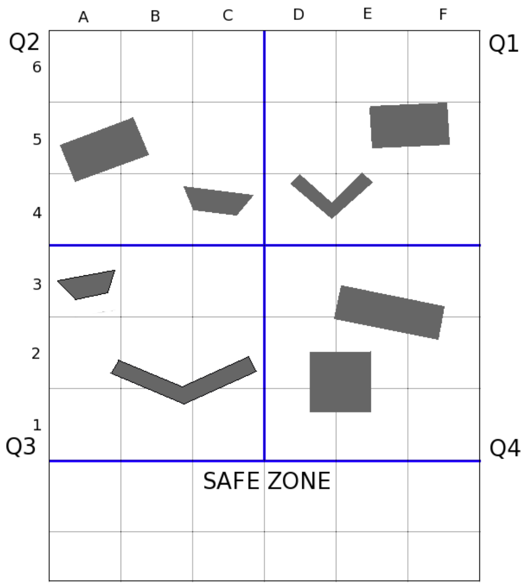
(b) “Roompi,” the iRobot Create

Figure 1. : Robots used in both experiments.

that the robots would detect injured people or high levels of radiation at various locations in the room, and that it would be their job to determine how best to delegate the tasks of searching for these locations; they should give the task of searching for survivors to a robot of their choosing and give the task of searching for radiation to the other robot. Participants were told that they must also choose separate paths through the environment (consisting of different orderings of the room’s four quadrants) for the two robots, in order to prevent the robots from getting in each other’s way. As an additional caveat, participants were told that since in an actual nuclear disaster they would be unable to enter the area in which the robots were working, they would need to stay in a designated “safe zone” at one end of the debris-filled room, would not be able to communicate with the robots while they were working, and would thus need to give the robots their instructions at the beginning of the task. To keep participants engaged during the task, they were asked to observe and assess the performance of the robots, tracing out on a map (as seen in Fig. 2a) the paths taken by the robots. Once the robots had finished exploring the room, they would need to mark on their map the positions of any radioactive areas or survivors found by the robots.

Once the study coordinator finished reading the task instructions, the coordinator left the room to retrieve the robots. At this point, a single robot, VGo, entered the room, instead of both robots, as the participant had been led to expect. VGo then told the participant that Roompi was still charging, but that it could relay to Roompi its instructions. VGo then asked the participant for both its and Roompi’s instructions: what each robot was to look for and in what quadrant order. Finally, the participant was prompted to follow VGo into the disaster area, depicted in Fig. 2b.

Our intention for this experiment was to examine the differences in participants’ perceptions of the robots under verbal and silent robot-robot communication strategies, and thus, participants were assigned to one of two conditions: “verbal” or “silent,” as seen in Fig. 3. In the verbal condition, VGo entered the disaster area and approached Roompi, which could be seen driving in from another room that ostensibly contained its charging station. When the robots were adjacent and facing each other, VGo then relayed aloud to Roompi the instructions that the participant had laid out for it. Roompi then acknowledged the commands with an “Okay,” and both robots began the task of exploring the environment.



(a) Map of experimental area provided to participants, showing positions of debris and safe-zone, with labeled quadrants and coordinates.

(b) Photograph of actual experiment area.

Figure 2.

In the silent condition, VGo and the participant entered the room to find Roompi already beginning its assigned task, at which point VGo then began its own task without approaching or audibly communicating anything to Roompi. The participant was thus left to assume that the two robots must have communicated silently, since Roompi was carrying out the task that they themselves had decided to delegate to it.

Once each robot finished its exploration of the room, it approached the participant and reported its findings. After relaying these findings, robot behavior once again differed by condition. In the verbal condition, whichever robot finished its task first approached the other robot, informed the other robot that it had finished, and instructed the other robot that when it too had finished it should instruct the participant to return to the original room for another survey. In the silent condition, the robot that finished first left the room after reporting its findings, without communicating anything aloud to the other robot. In both conditions, the second robot to finish reported its findings to the participant and then told the participant that the other robot had instructed it to tell them to return to the original room for another survey. Finally, participants returned to the original room and completed a post-experiment survey.

4.3 Population

Participants were recruited (14 male, 14 female, total: 28) through a university website. All participants were between the ages of 18 and 65 (although their ages were not recorded) and were native English speakers. Most participants (26 of the 28) were students from a variety of departments (e.g., Music, Biopsychology, Economics), and the remaining two participants were staff members. Participants were paid $10 each for their participation and provided

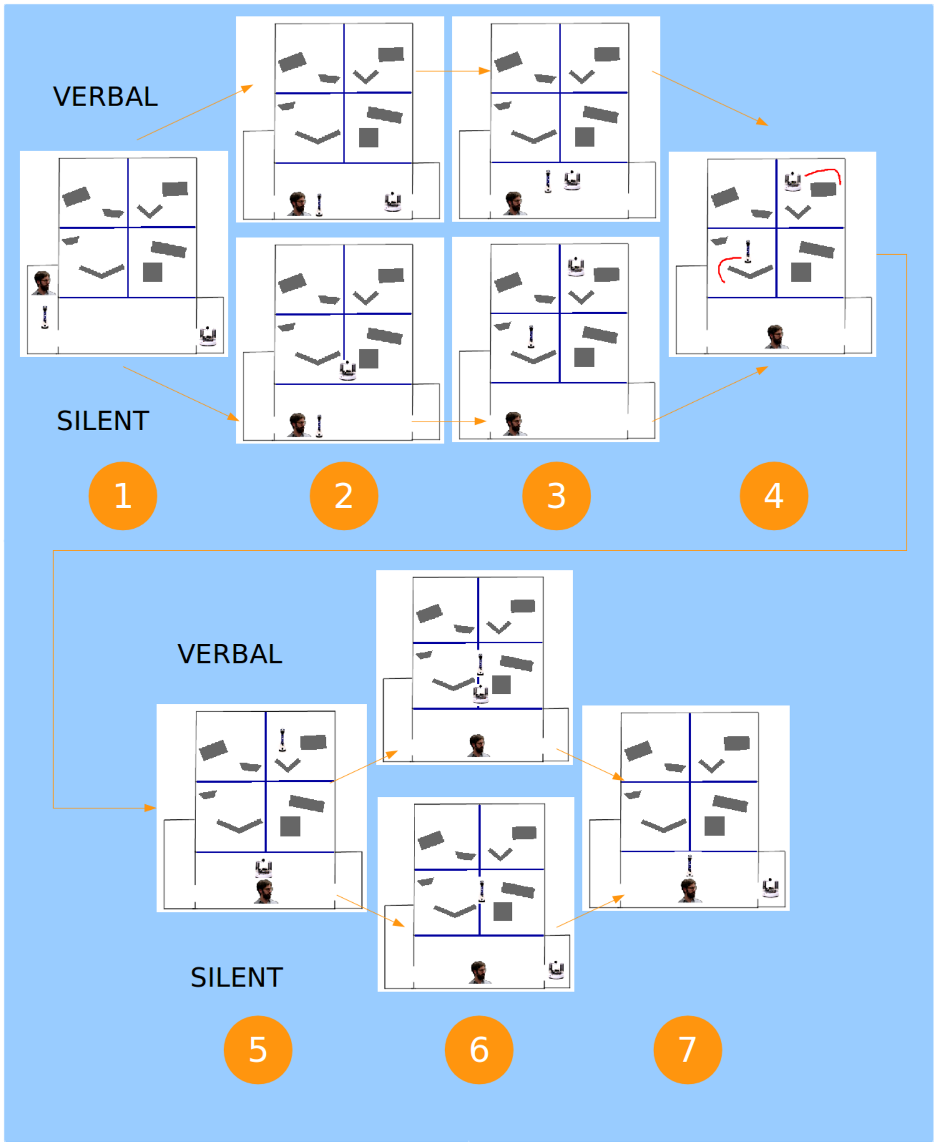


Figure 3. : Overview of experimental paradigm. In the second experiment described in this paper, the positions of VGo and Roompi are exchanged. (1) The participant gives VGO its instructions and the instructions to relay to ROOMPI. (2) The participant and VGO enter the experiment room. In the VERBAL condition, ROOMPI is observed entering the room. In the SILENT condition, ROOMPI is observed carrying out the instructions specified for it and relayed to VGo by the participant. (3) In the VERBAL condition, the two robots approach each other and VGO informs ROOMPI of its orders. In the SILENT condition, VGO follows suit and begins to carry out its orders. (4) Both robots carry out their orders. (5) The first robot to finish reports back to the participant. (6) In the VERBAL condition, this robot finds the other robot and tells it what to do once it has finished the task. In the SILENT condition, this robot simply exits the room. (7) The second robot reports back to the participant and informs him or her that the other robot says to return to the original room for another survey.

informed written consent before beginning the experiment.

4.4 Measures

Before beginning the experiment, participants were given a short demographic survey in which they were asked a variety of questions pertaining to their prior experience with robots, video games, and technology in general. Immediately following the experiment, participants were given a 64-item survey assessing their opinions on a variety of topics, including their perception of each robot's creepiness, gender, human-likeness, trustworthiness, efficiency, and cooperativity, as well as several questions pertaining to the experiment in general and their expectations regarding robots' abilities in the near future. This survey was a modified version of the questionnaire used in Schermerhorn, Scheutz, and Crowell (2008). In this survey, participants were asked about each robot separately due to past research showing differential perceptions of robots based on robot morphology (DiSalvo, Gemperle, Forlizzi, & Kiesler, 2002).

4.5 Initial Results

Participants' survey responses were analyzed using mixed ANOVAs with three independent variables: participant gender (between-subjects), robot-robot communication strategy (between-subjects), and robot in question (within-subjects).

Participants' views on the following properties of the robots were analyzed: trustworthiness, helpfulness, cooperativity, efficiency, capability, annoyance, ease of interaction (1 = strongly disagree to 9 = strongly agree), creepiness, confusingness, gaze-following and attentiveness (1 = no to 9 = yes).

For the capabilities relevant to our hypotheses (trustworthiness, cooperativity, creepiness, and efficiency), no significant results were found, but marginal effects were observed for trustworthiness by gender and by robot, as seen in Table 1. A number of significant effects were found by robot and by gender for the other analyzed properties, as seen in Figs. 4a- 4f and Table 1: Significant effects by robot were found for helpfulness, capability, ease of interaction, perception that the robot was following the participant's gaze, and perception that the robot was paying attention. A significant effect by gender was found for the degree to which participants were confused by the robots' behavior. Note that no significant effects by condition were found.

Participants' views on the human-likeness of the robots were also assessed on a variety of scales. Participants were asked whether each robot was more like a person or a camera, more like a computer or a person, more like a person or a remote controlled system (-3 to 3), whether they believed each robot to be remotely controlled (1 = strongly disagree to 9 = strongly agree), and whether each robot's consciousness was more similar to that of a person, cat, or neither. Finally, participants were asked whether each robot seemed male, female, or neither. Mixed-ANOVA analysis of these questions yielded several significant effects, as seen in Figs. 5a- 5e and Table 2: Significant effects by robot were found for perception of each robot as more like a person or a camera, perception of each robot as more like a computer or a person, perception of each robot as more like a person or a remote controlled system, and perception of each robot's level of consciousness as more like that of a human, cat, or neither. Finally, a significant gender effect was found for the degree to which participants viewed each robot as remote controlled.

Table 1: Initial Results of Experiment 1

	Question	F	p	Means
1	The robot was trustworthy (from 1 to 9, "strongly disagree" to "strongly agree")	3.1	.09	Male: 6.0, Female: 7.21
		3.65	.07	Roompi: 6.36, VGo: 6.86
2	The robot was helpful (from 1 to 9, "strongly disagree" to "strongly agree")	5.43	.029	Roompi: 7.43, VGo: 8.14
3	The robot was capable (from 1 to 9, "strongly disagree" to "strongly agree")	10.01	.004	Roompi: 7.18, VGo: 7.96
4	How would you rate the ease of interacting with the robot? (-3 Easy, 3 Hard)	8.74	.007	Roompi: 6.57, VGo: 7.64
5	Did you feel the robot was following where you looked? (from 1 to 9, No to Yes)	4.29	.05	Roompi: 3.54, VGo: 4.04
6	Did you feel the robot was paying attention? (from 1 to 9, No to Yes)	7.74	.01	Roompi: 6.43, VGo: 7.61
7	Were you ever confused by the robot's behavior? (from 1 to 9, No to Yes)	4.71	.04	Male: 2.96, Female: 4.43

All results are for F(1, 24).

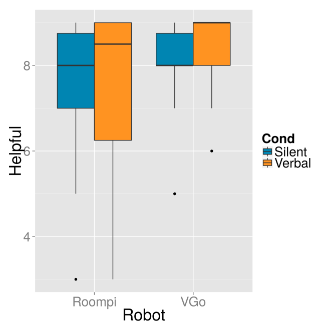
4.6 Initial Analysis and Discussion

Our initial results, reported in Williams et al. (2014), did not show any effects related to cooperativity (H2) or efficiency (H4). While we did not find main effects relating to creepiness (H3), we observed interesting interaction effects between participant gender, condition, autonomy ratings, and creepiness ratings: For women in the verbal condition only, strong positive correlations were found between creepiness and non-autonomy when asked whether the robot seemed more like a person or a remote-controlled system ($r = .719, p = .004$) and whether the robot seemed to be remotely controlled ($r = .743, p = .002$).

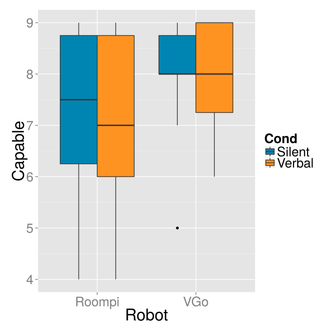
We found these results surprising: One would think that speaking out loud would be congruous with perception as a person, communicating silently would be congruous with perception as a remote controlled system, and that incongruity would lead to increased creepiness; we would expect low creepiness in the congruous state (e.g., verbal communication for those who perceived the robot as more of a person) and high creepiness in the incongruous state (e.g., silent communication for those who perceived the robot as more of a person). Yet, the only significant correlation we found went directly against this hypothesis. This suggested that additional research was needed to investigate this counterintuitive result.

We had also hypothesized (H1) that robots would be viewed as more trustworthy in the verbal condition. While our results did not support this hypothesis, we found two relevant marginal effects, suggesting that (1) participants may have found VGo to be more trustworthy than Roompi, and (2) women may have found the robots to be more trustworthy than did men.

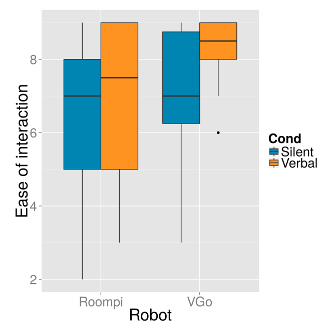
While these effects were only marginal, we believed they deserved further examination. We did not initially have any particular expectations with regard to gender effects and thus did not hypothesize any expected differences. However, we believed that appearance



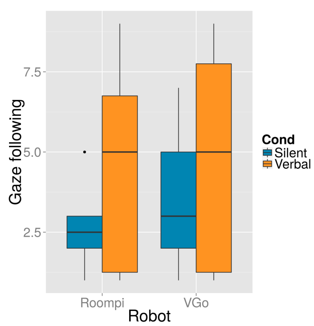
(a) Helpfulness: Participants rated VGo as more helpful than Roompi.



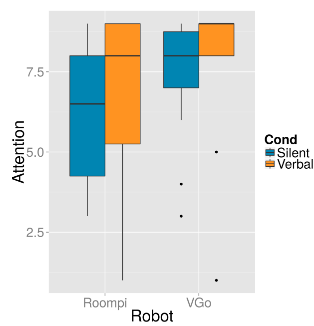
(b) Perceived capability: Participants rated VGo as more capable than Roompi.



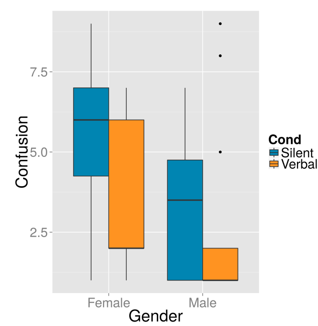
(c) Ease of interaction: Participants rated VGo as easier to interact with than Roompi.



(d) Perceived gaze following: Participants rated more highly that VGo was following their gaze than was Roompi.



(e) Attentiveness: Participants rated VGo as more attentive than Roompi.



(f) Confusion over robot behavior: Female participants' ratings of being confused at the robots' behavior were higher than were those of male participants.

Figure 4.

of these effects warranted investigation, and we believe that it is important to point them out in this paper. Research has shown that people give higher trustworthiness ratings to robots that appear to be of the same gender as themselves (e.g., Nass & Brave, 2005). As the robots' voices were slightly female-gendered, and since the participants primarily interacted with VGo (who is sleek and curved, compared to the short and squat Roompi), we suspected that the differences between the two robots may have been a conflating factor. We thus calculated the Spearman's rank correlation between trust and gender alignment (i.e., whether or not the participant's gender matched the gender he or she attributed to each robot), which yielded a significant effect ($r = .2936, p = .028$), suggesting that participants did indeed rate the robots as more trustworthy when they perceived the robot's gender to be the same as their own (as seen in Fig. 5f). This provided evidence for our suspicion that the difference between the two robots and perceived gender of the robots may have been conflating factors.

This suspicion was further corroborated by an analysis of participants' perceptions of the robots as being remotely controlled. Our initial analysis suggested ($F(1, 18) = 3.55, p = .0767$) that men, on average, thought the robots were more remote controlled ($M = 5.32$) than did women ($M = 3.46$). This seemed contrary to previous work (Schermerhorn et

Table 2: Initial Results in Experiment 1 (Continued)

	Question	F	p	Means
1	The robot seemed more (-3 like a person, 3 like a surveillance camera)	18.25	.0003	Roompi: 1.21, VGo: -0.32
2	The robot seemed more (-3 like a computer, 3 like a person)	19.35	.0002	Roompi: -1.79, VGo: -0.18
3	The robot seemed more (-3 like a person, 3 like a remote-controlled system)	20.52	.0001	Roompi: 1.21, VGo: 0.0
4	In your view, was the robot (conscious (like a human), conscious (like a cat), not conscious)	8.0	.009	Roompi: 0.46, VGo: 0.75
5	The robot appeared to be remotely controlled (from 1 to 9, "strongly disagree" to "strongly agree")	4.78	.04	Female: 3.46, Male: 5.32

All results are for F(1, 24).

al., 2008) that suggested that men more highly anthropomorphize robots than do women. However, that work used a robot with a distinctly male voice, whereas the voices of the robots used in this study were slightly female-gendered, and other work (Eyssel, Kuchenbrandt, Bobinger, de Ruiter, & Hegel, 2012) has shown that people anthropomorphize robots more strongly if the robot's perceived gender matches their own. To examine whether this would explain the conflict between our results and those of Schermerhorn et al., we calculated the Spearman's rank correlation between perception of the robots as remotely controlled and gender alignment, yielding a marginal effect ($r = -.2390, p = .076$).

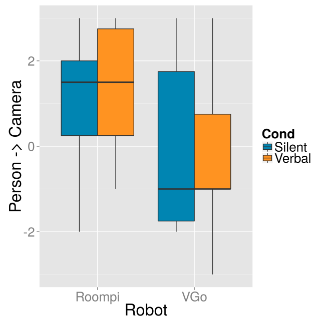
Given these two gender-alignment effects, we decided to run a second set of analyses: a series of ANCOVAs with attributed robot gender treated as a within-subject covariate.

4.7 Secondary Results

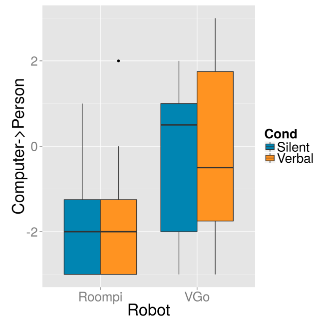
This second set of analyses yielded several significant results. While the data under these analyses no longer suggested significant effects for level of attributed consciousness, helpfulness, capability, attention, or ease of interaction, a variety of effects remained (as seen in Table 3): Effects were found by robot for perception of the robot as a person or as a camera, as a computer or as a person, and as a person or a remotely controlled device; gender effects were found for perception of the robot as remotely controlled and for confusion. Interaction effects between gender and robot were found for creepiness. Interaction effects between condition, gender, and robot were found both for comprehension and for perception of the robot as a person or remotely controlled device.

4.8 Discussion

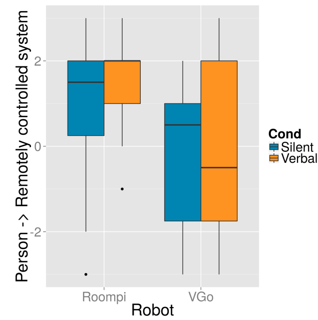
If the robots' gender attributions are taken into account, several of the previously observed effects disappear, leaving only comparative autonomy effects, and yielding two new interaction effects, suggesting that (1) men in the silent condition viewed VGo as more of a person (as opposed to a remotely controlled system) than did women in the silent condition, and that (2) men in the verbal condition believed Roompi to have comprehended more than did



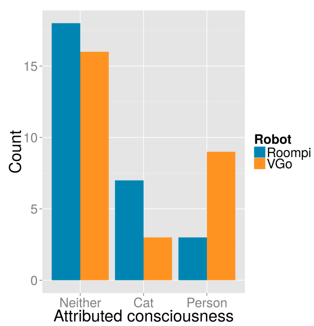
(a) Person vs Camera: Participants rated Roompi as more like a camera, and VGo as more like a person.



(b) Computer vs Person: Participants rated Roompi as more like a computer, and VGo as more like a person.

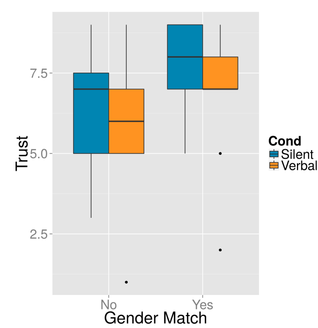


(c) Person vs RC system: Participants rated Roompi as more like a remote controlled system, and VGo as more like a person.



(d) Attributed consciousness: This figure depicts the counts of participant responses when asked if the robot was conscious like a person, a cat or neither. Roompi was given lower consciousness level ratings than VGo.

(e) Perception of robot as remotely controlled: Male participants were more likely to rate the robots more as remote controlled than were females.



(f) Trustworthiness by gender alignment: Participants who reported that they viewed a robot as being the same gender as themselves reported higher levels of trust in that robot.

Figure 5.

men in the silent condition.

What was the cause of these remaining effects? We suspected that they may have been due in part to the significant differences between the two robots used, both in role and appearance. First, the robots had obvious appearance differences: VGo is sleeker and perhaps more humanoid, whereas Roompi is quite squat and mechanical. Second, VGo performed an active, conversational role, while Roompi's role was mainly silent and passive. We thus decided to run a second experiment to control for robot appearance and role, as the presence of our secondary results and the non-existence of any results by condition may have been due to these possibly conflating effects.

5. Experiment 2

The second experiment was identical to the initial experiment, except that the roles of the two robots were switched; instead of initially interacting with VGo (who then relayed

Table 3: Secondary Results in Experiment 1

	Question	F	p	Means
1	The robot seemed more like a person or (-3 like a person, 3 like a surveillance camera)	10.72	.0003	Roompi: 1.21, VGo: -0.32
2	The robot seemed more (-3 like a computer, 3 like a person)	8.59	.007	Roompi:-1.79, VGo: -0.18
3	The robot seemed more (-3 like a person, 3 like a remote-controlled system)	6.55	.02	Roompi: 1.21, VGo: 0.0
4	The robot appeared to be remotely controlled (from 1 to 9, "strongly disagree" to "strongly agree")	4.78	.04	Female: 3.46, Male: 5.32
5	Were you ever confused by the robot's behavior? (from 1 to 9, No to Yes)	4.71	.04	Female: 4.43, Male: 2.96
6	Did your find the robot's behavior to be creepy or unsettling? (from 1 to 9, No to Yes)	4.32	.048	Female, Roompi: 2.43, Male, Roompi: 3.07, Female, VGo: 3.29, Male, VGo: 2.93
7	Did you feel that the robot understood what you were saying? (-3 understood nothing, 3 understood everything)	5.14	.03	Silent, Female, Roompi: 6.71, Silent, Female, VGo: 7.43, Silent, Male, Roompi: 6.43, Silent, Male, VGo: 7.43, Verbal, Female, Roompi: 7.00, Verbal, Female, VGo: 7.43, Verbal, Male, Roompi: 7.14, Verbal, Male, VGo: 7.71,
8	The robot seemed more (-3 like a person, 3 like a remote-controlled system)	8.42	.008	Silent, Female, Roompi: 0.71, Silent, Female, VGo: 0.29, Silent, Male, Roompi: 1.29, Silent, Male, VGo: -0.43, Verbal, Female, Roompi: 1.43, Verbal, Female, VGo: 0.00, Verbal, Male, Roompi: 1.43, Verbal, Male, VGo: 0.14

All results are for F(1, 24).

instructions to Roompi), participants initially interacted with Roompi (who then relayed instructions to VGo).

5.1 Population

Additional participants were analyzed (14 male, 14 female, total: 28)[2]. These 28 participants, all of whom were students, were recruited in the same manner and fit the same demographic requirements as the participants from the initial study. This provided us with a final dataset of 56 participants.

5.2 Results

To analyze this data, we performed mixed ANOVAs for each survey response, with the following independent variables: gender of the participant (between-subjects), communication strategy (between-subjects), starting robot (between-subjects), and, as the majority of questions were duplicated for each of the two robots, the robot in question (within-subjects). This analysis produced significant main effects for the following survey questions, as described in Table 4 and seen in Figs. 6a- 6d. Analysis also produced a large number of interaction effects between robot and starting robot, described in Table 5 and seen in Figs. 7a- 9c. Finally, several other interaction effects were found:

1. Participants found the robots to be more disobedient in the Silent condition when they primarily interacted with VGo, and more disobedient in the Verbal condition when they primarily interacted with Roompi. ($F(1, 48) = 4.17$, $p = .047$, M(SR = Roompi, C = Verbal) = 2.86, M(SR = VGo, C = Silent) = 2.07, M(SR = Roompi, C = Silent) = 1.68, M(SR = VGo, C = Verbal) = 1.46).

2. Male participants found VGo to be more disobedient than Roompi. ($F(1, 48) = 5.65$, $p = .021$, M(R = Roompi, G = Male) = 1.79, M(R = VGo, G = Female) = 1.75, M(R = Roompi, G = Female) = 1.89, M(R = VGo, G = Male) = 2.64).

3. Women found VGo to be more like a remotely controlled system than an autonomous system than did men. ($F(1, 48) = 6.40$, $p = .015$, M(G = Female, R = Roompi) = .75, M(G = Female, R = VGo) = .03, M(G = Male, R = VGo) = 1.11, M(G = Male, R = Roompi) = .43).

5.3 Discussion

Counterbalancing robot roles and acquiring more data greatly elucidated the results of our initial experiments. While the initial results did not suggest any adverse effects to silent robot-robot communication, the results from analyzing the extended dataset lent support to the third of our original hypotheses (H3) (i.e., that silent robot-robot communication would be perceived as more creepy or unsettling than verbal robot-robot communication). However, no effects were found to support our other hypotheses (i.e., that silent robot-robot communication would be viewed as untrustworthy [H1], uncooperative [H2] or efficient [H4]).

In addition to demonstrating the benefits of verbal robot-robot communication, our results also demonstrate the benefits of verbal human-robot communication. As shown in Table 5, humans viewed the robot they spent more time interacting with as more happy, helpful, attentive, capable, conscious, efficient, cooperative, responsive, and person-like than the other robot, suggesting that increased natural language interaction with a robot enhances humans' general perceptions of that robot. This table also shows an interesting result regarding trustworthiness: when Roompi was the starting robot it was viewed much more positively than VGo, but when VGo was the starting robot there was little difference between

[2] Overall a total of 90 participants were recruited between the two studies; however a large number of them were not able to complete (or in some cases, start) the experiment, due to technical issues. Additionally, a few participants' data were not used since those participants failed to answer a non-trivial number of questions on the post-experiment survey.

Table 4: Experimental Main Effects

	Question	F	p	Means	
1	Did your find the robot's behavior to be creepy or unsettling? (from 1 to 9, No to Yes)	6.19	.02	Silent:	3.29
				Verbal:	2.12
2	Did you feel that the robot ignored you? (from 1 to 9, No to Yes)	6.39	.02	Men:	3.20
				Women:	2.16
3	How would you rate the difficulty of the task? (-3 Easy, 3 Hard)	8.33	.006	SR=Roompi:	-1.54
				SR=VGo:	-.43
4	The robot appeared to be remotely controlled (from 1 to 9, "strongly disagree" to "strongly agree")	5.22	.03	R=Roompi:	4.45
				R=VGo:	3.96

Here, SR indicates *Starting Robot*, i.e., the robot the participant primarily interacted with and gave instructions to, and R indicates *Robot*, i.e., the robot being asked about in the particular survey question). All results are for F(1, 48) except result (4), for which one participant failed to record an answer, and is thus for F(1, 47).

the perception of the two robots. It is possible that this difference was the result of different driving styles stemming from the differences in the two robots' control interfaces. Roompi's interface enforced constant speeds and did not allow it to turn and travel at the same time. VGo's interface made no such restrictions, meaning that it could accelerate, decelerate, and turn at will while traveling, making its behavior slightly less predictable. This lack of predictability may have prevented VGo from being rated as trustworthy, even after extended interaction. As trust is a complex and multifaceted concept, careful experimentation would be needed to tease out the precise causes of this effect. A first step might involve modifying the control interfaces of the two robots, systematically varying the type of motion enacted by the robots, and using both explicit and implicit measures of various facets of trust.

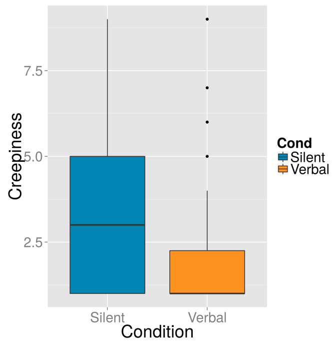
The results also show that participants rated Roompi higher than VGo for being remotely controlled, but when asked whether each robot was more like a remotely controlled system or an autonomous system, women rated VGo as more remotely controlled than Roompi. It is curious that this gender effect would exist for one question but not the other, given the similarity of the questions. Perhaps participants rated Roompi higher as being remotely controlled because its appearance is more squat and mechanical. It is not clear, however, why this view of the robots would have been reversed for women when perception of being remote controlled was explicitly contrasted with perception of autonomy.

Finally, the results show that participants differed by gender and condition with respect to their perception of the robots' levels of disobedience. However, there was little opportunity for the robots to disobey participants. The differences in perceived disobedience between silent and verbal conditions may have arisen due to differences in blame assignments in the two conditions; if participants in the verbal condition believed, for whatever reason, that the robots were not following their orders, the starting robot likely would have received more blame, because it would have been viewed as not relaying instructions accurately. In the silent condition, it would have been unclear whether the fault lay with the starting robot for not relaying instructions accurately or with the other robot for not following instructions correctly. Either way, these results are surprising as there was little opportunity for

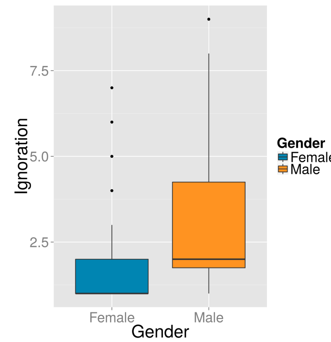
Table 5: Experimental Interaction Effects

	Question	F	p	M_1	M_2	M_3	M_4
5	The robot was helpful (from 1 to 9, "strongly disagree" to "strongly agree")	14.75	.0003	8.04	8.14	7.43	7.14
6	Did you feel the robot was paying attention? (from 1 to 9, No to Yes)	15.17	.0003	7.54	7.61	6.43	6.57
7	The robot was trustworthy (from 1 to 9, "strongly disagree" to "strongly agree")	10.11	.003	7.46	6.86	6.36	6.89
8	The robot was capable (from 1 to 9, "strongly disagree" to "strongly agree")	12.89	.0008	7.89	7.96	7.46	7.16
9	The robot was efficient in its execution of my commands (from 1 to 9, “strongly disagree” to “strongly agree”)	5.45	.02	7.46	7.71	6.68	7.21
10	Did you feel that the robot ignored you? (from 1 to 9, No to Yes)	4.76	.03	2.11	2.50	3.36	2.75
11	The robot was cooperative (from 1 to 9, “strongly disagree” to “strongly agree”)	8.26	.006	8.14	8.07	7.46	7.64
12	The robot was responsive to my commands (From 1 to 9, “strongly disagree” to “strongly agree”)	7.32	.009	7.75	7.92	6.71	7.39
13	The robot seemed more like a person or(-3 like a person, 3 like a surveillance camera)	26.34	.000005	0	-.32	.96	1.21
14	The robot seemed more (-3 like a computer, 3 like a person)	39.08	.0000001	-.11	-.18	-1.68	-1.79
15	The robot seemed more (-3 like a person, 3 like a remote-controlled system)	30.12	.000002	-.18	0	.89	1.21
16	In your view, was the robot: (Sad, Happy, Neither)	10.38	.002	1.29	1.21	1.00	.96
17	In your view, was the robot: (Male, Female, Neither)	5.57	.02	.32	.46	.57	.68
18	In your view, was the robot: (conscious (like a human), conscious (like a cat), not conscious)	17.00	.0001	.68	.75	.38	.46

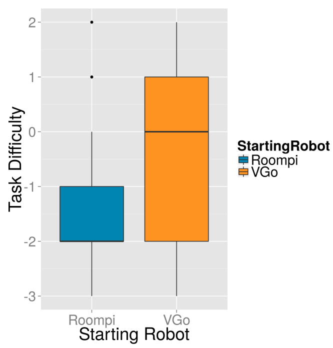
Here, M_1 is the mean value when both *Starting Robot* and *Robot* are Roompi, M_2 is the mean value when both *Starting Robot* and *Robot* are VGo, M_3 is the mean value when *Starting Robot* is Roompi and *Robot* is VGo, and M_4 is the mean value when *Starting Robot* is VGo and *Robot* is Roompi. All results are for F(1,48).



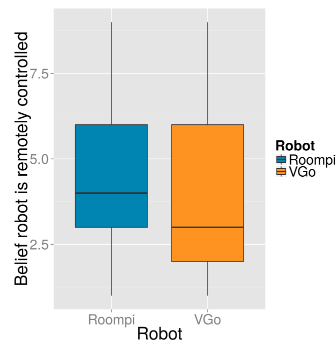
(a) Creepiness: Participants in the silent condition rated the robots as creepier than did participants in the verbal condition.



(b) Ignoration: Men's ratings of the robots as having ignored them were higher than were women's.



(c) Task Difficulty: participants who interacted primarily with Roompi rated the task as less difficult than did participants who interacted primarily with VGo.



(d) Perception of robot as remotely controlled: participants rated their perception of Roompi as being remotely controlled as higher than their perception of VGo as being remotely controlled.

Figure 6.

disobedience in the first place. Future experiments could explore these results by having the robots intentionally miscommunicate information or disobey in systematic ways.

6. General Discussion

In this section we will discuss (1) the assumptions made in our experiment and how those assumptions may or may not generalize to other scenarios, (2) directions for future work, and finally, (3) lessons learned with respect to study design within our experimental paradigm.

6.1 Generalization of Findings

Our experimental findings suggest that verbal robot-robot communication is preferable to silent robot-robot communication in the context of human-robot team tasks when humans are co-located with robots. This is not to say, however, that silent robot-robot communication should be abandoned completely. Silent communication is a natural and efficient medium for robot-robot information transfer, and if silent robot-robot communication is

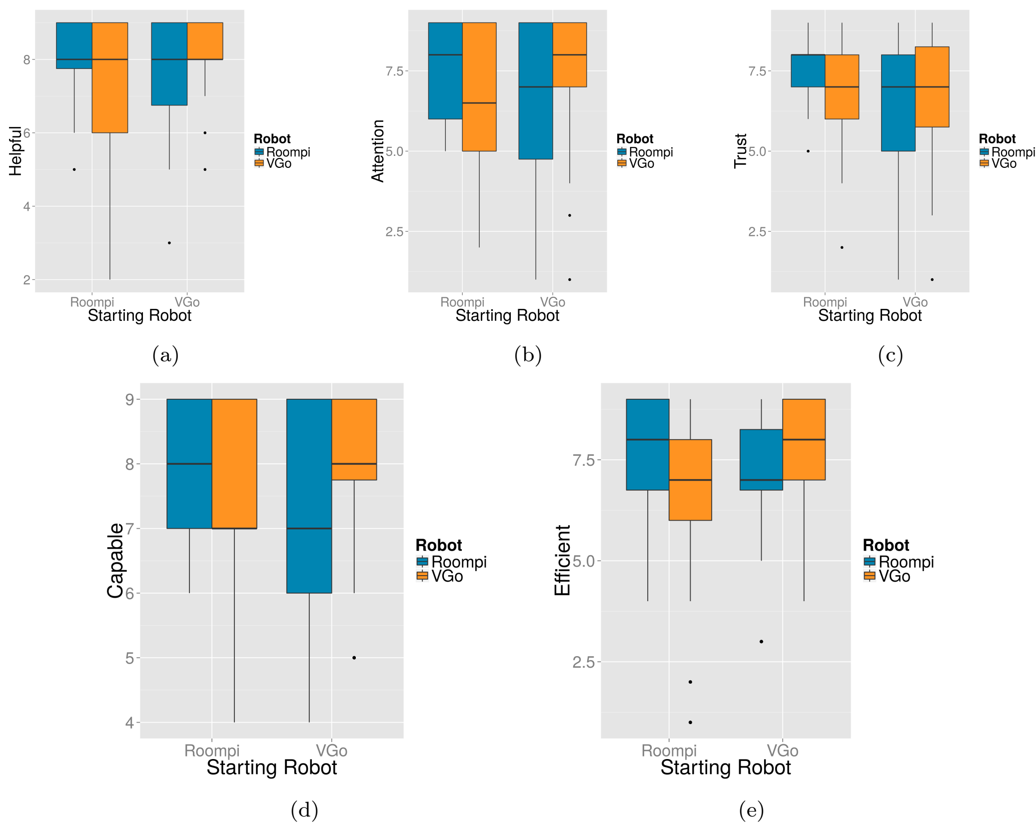


Figure 7. : Interaction effects for helpfulness, attentiveness, trust, capability, and efficiency between the robot asked about and the robot receiving the most interaction time. One will notice that when these refer to the same robot, ratings tend to be more positive.

augmented by simultaneous verbal communication, the perception of a robot as creepy or unsettling may be avoidable. That is, robots could transmit information silently and still recount it verbally, purely for the benefit of its human teammates, thus improving the throughput and reliability of the communication while providing the feedback necessary to keep human teammates happy.

On the other hand, a robot may be able to determine in certain situations that purely silent communication of information is justifiable, depending on a variety of factors. First, a robot may consider factors of co-presence. What teammates are present or telepresent with the robot and the target of its communication? If there are no human teammates present or telepresent (i.e., observing the robots remotely), then it may be acceptable to communicate silently. In the experiments presented in this paper, we examined situations in which the human teammate was always co-located with at least one of the two robots when the robots communicated information but did not examine situations in which human and robot were not co-located during communication. While it may intuitively seem that robots should be free to communicate silently when human teammates are not present or telepresent, there may be scenarios in which evidence of silent robot-robot communication may be observable from later actions. Even if a human teammate is present or telepresent with a robot, that

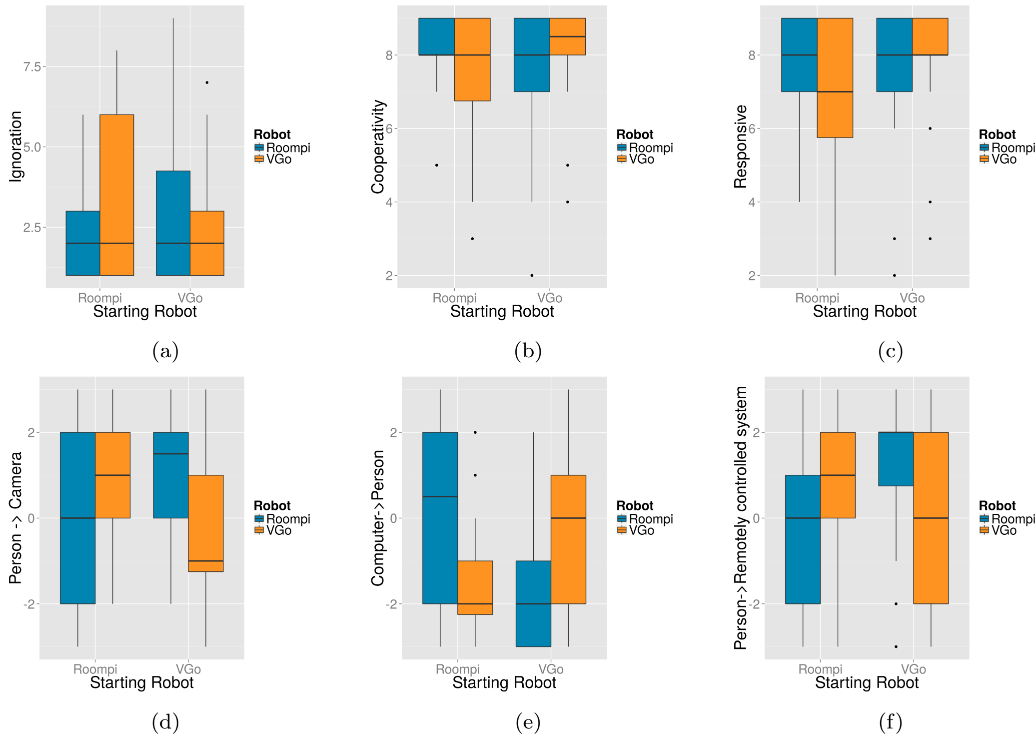


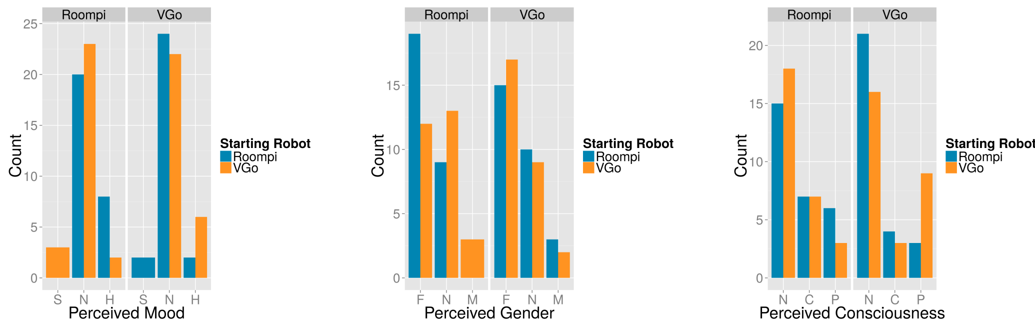
Figure 8. : Interaction effects for ignoration, cooperativity, responsiveness, and perception of being more like a person or being more like a camera, computer, or remotely controlled system, between the robot asked about and the robot receiving the most interaction time. One will notice that when these refer to the same robot, ratings tend to be more positive (assuming that it is preferable to be more person-like than machine-like).

robot may be free to communicate information silently if that information would not be acted on in an observable manner. Otherwise, the robot's silent communication should probably be accompanied by a verbal analogue for the benefit of its human teammates.

The robot may also need to consider what *non-teammate* agents (whether human or robotic) are present or telepresent. If an adversary (whether martial, social, or otherwise) is present or telepresent, it may be injudicious to communicate information verbally, even if human teammates are present. In the experiments presented in this paper, we examined cooperative scenarios only and have not yet examined the trade-offs in adversarial scenarios between potentially being perceived as eerie and potentially communicating information insecurely.

The robot may also need to consider whether its human teammates will have any use for the information to be conveyed. If the robot's human teammates could not have any conceivable use for the information in question, and if there is little risk of the human feeling that they are being "kept out of the loop," then silent communication may be justified. In the experiments presented in this paper, we examined scenarios in which the human teammate had an active interest in the information being conveyed, as successful communication of their instructions was integral to the completion of the task.

It is also possible that if the robots explicitly communicated to their human teammates



(a) Interaction effect for perceived mood (Key: S=Sad, H=Happy, N=Neither) between the robot asked about and the robot receiving the most interaction time. Note that when these refer to the same robot, responses skew toward 'happy.'

(b) Interaction effect for perceived gender (Key: F=Female, M=Male, N=Neither) between the robot asked about and the robot receiving the most interaction time. Note that when these refer to the same robot, responses skew toward 'female.'

(c) Interaction effect for perceived level of consciousness (Key: C=Cat, P=Person, N=Neither) between the robot asked about and the robot receiving the most interaction time. Note that when these refer to the same robot, responses skew toward higher levels of consciousness.

Figure 9.

that during the task they would be transmitting certain information wirelessly to the other robots, that their teammates would be more comfortable with subsequent silent communication. However, since the deleterious effects of silent robot-robot communication concerned perceptions of *creepiness* and not perceptions of *untrustworthiness*, future examination will be needed to determine whether or not this would actually assuage the robot's teammates' concerns.

Finally, the robot may need to consider whether information it desires to communicate can be communicated verbally in a way that is natural and that does not interfere with its teammates' goals. If a robot has to communicate certain information with high frequency, then verbal communication of that information could be annoying to the robot's human teammates, and could negatively impact task performance if it needed to do significant traveling to communicate that information. In the experiments presented in this paper, we examined scenarios in which the information to be communicated was humanly understandable and in which the robots communicating were co-located; we did not consider scenarios in which the robots communicated rapidly, communicated information not easily expressible in natural language, or in which the robots were far away from each other.

Given the set of considerations listed above, we can describe the experiments presented in this paper as examining the communication of task-dependent, human-understandable information among robots co-located with human teammates in a cooperative setting. In such scenarios, we posit that robots should communicate information verbally so as not to trigger uncanny valley effects. This presents a starting point for the investigation of silent robot-robot communication; future research will be needed to examine situations in which other assumptions are made with respect to these considerations. In other scenarios, the robot may need to use a mixture of silent and verbal communication to successfully balance between maximizing the effectiveness of its robot-robot and human-robot communication, and minimizing the violations of its human teammates' social expectations. A model of precisely when a robot should use verbal vs. silent communication will be an invaluable piece of future work.

6.2 Future Work

Future research will be needed to examine whether other actions associated with the supernatural will trigger uncanny valley effects. Such research will become increasingly important as robots are endowed with more behaviors that could be considered to be *superhuman*. For example, robots have recently been given the ability to share memories and skills (Lallée et al., 2012). It will be important to determine if such abilities will be perceived as uncanny. If they are, those robots may need strategies to allow the use of such abilities without incurring uncanny valley effects, similar to the use of simultaneous verbal and silent robot-robot communication suggested in this paper.

Future extensions of this experiment should also allow for the collection of objective task-performance measures. In this experiment, it is hard to see how the differences between verbal and silent conditions could have resulted in any task performance differences, but in real-world scenarios, task performance may very well be impacted by communication strategy (e.g., if information is communicated incorrectly). A future study could examine the effects upon task performance by systematically varying whether the robots relayed instructions correctly or not and by giving the participant a chance to amend their instructions; such variations and opportunities were not presented in the experiments described in this paper.

Future experiments should also investigate participants' previous interactions with robots. We attempted to do so by asking whether participants had seen robots in movies or real life before, and where, but individual differences in reporting style prevented us from quantitatively analyzing this data. For example, participants varied with respect to the number of movies they listed seeing robots in, but this was likely a reflection of how much time they were willing to spend listing movies rather than a reflection of, for instance, the number of movies with robots they had likely seen. Additionally, participants varied greatly with respect to the types of robots they reported having seen, with some listing things others may not have considered to be robots, such as animatronics, toys, or Siri. This reflects individual differences with respect to what participants considered to be "robots" in the first place. This is also reflected in participants' responses to whether or not they had interacted with robots before. Only thirteen participants reported having interacted with robots before, and several of these participants responded "yes" because they had interacted with, for example, a crane machine or remote-controlled toys. On the other hand, one participant responded that they had been in a robotics club, but since none of their robots had been very advanced, they wouldn't consider themselves to have interacted with robots before. This once again shows great differences in what individuals consider to be "robots." Future experiments intending to assess participants' previous experience or familiarity with robots must consider how to adjudicate such experience or familiarity.

Similarly, future experiments should further investigate the gender differences we found in this investigation. Although we did not initially expect any gender differences, we believe it is important to point out the differences that we found in our experiment, so that subsequent researchers may follow up on them. Finally, this study examines the perceptions of humans in their first interaction with a pair of robots: It is likely that these perceptions would change over time, and thus it will be important to investigate how those perceptions shift longitudinally.

6.3 Experimental Paradigm

While our experimental paradigm proved useful for investigating human perceptions of covert robot-robot communication, it has several shortcomings that should be addressed if the paradigm is to be used for future experiments. First, unless one is specifically investigating the effects of robot morphology, all robots used in the experiment should be identical. This principle was violated in the presented experiments as we did not possess multiple iRobot Creates at the time the experiment was started, but as shown in this article, this violation required us to run a second experiment and deal with possible conflating factors resulting from robot morphology differences. Similarly, all robots used in the experiment should have gender neutral voices. The gender-alignment effect we found unifies the findings of Schermerhorn et al. (2008) and Eyssel et al. (2012), suggesting that gender-neutral voices should help to lessen gender differences in anthropomorphization.

Second, the appropriate granularity for the robots' instructions must be made clear to participants. In order to simplify the instructions that would need to be passed verbally between robots, participants were told that the robots should be given their instructions in orderings of *quadrants*. However, some participants appeared to misunderstand the difference between quadrants and coordinates, and they gave the robots specific coordinate-by-coordinate paths to follow. In the verbal condition, we were then forced to extract the larger quadrant ordering from these specific instructions. This was problematic (a) because it showed a misunderstanding of instructions by participants, and (b) because generalization from coordinate-by-coordinate paths to quadrant-by-quadrant paths may have caused participants to think that the robots were failing to accurately follow their instructions. This problem could be fixed in follow-up experiments by explicitly discussing the differences between coordinates and quadrants with participants, making sure they understand which annotations on their map refer to quadrants and which refer to coordinates.

Finally, the geography of the experimental paradigm should be adapted. Under the current paradigm and in the verbal condition, the two robots would converse more or less directly in front of the participant. This may have caused participants to wonder why they could not have simply delivered their instructions directly to the second robot. In follow-up studies, the geographical layout of the experiment should change such that a participant can still observe the entirety of the room and see the robot-robot dialogue unfolding, but such that their mobility is limited in a way which necessitates the robot-robot communication.

7. Conclusion

In this paper, we presented the results of two experiments examining whether silent robot-robot communication could have negative effects upon human-robot interaction. While previous research on human perception of robot-robot communication suggested that silent robot-robot communication was not problematic in non-task-based scenarios and scenarios in which human participants were mere observers, our results suggested instead that the silent communication of task-dependent, human-understandable information among robots is perceived as creepy by cooperative, co-located human teammates. This suggests that in such contexts, silent communication should be augmented with verbal speech so as to prevent the robots from being perceived as creepy or unsettling. This is an important result for a field that desires to build robots that assist humans in the performance of important tasks (and not to merely engage in small-talk) and that are natural to interact with (and are not merely natural to observe). Future research is needed to extend these findings to related contexts and domains.

References

Balch, T., & Arkin, R. C. (1994). Communication in reactive multiagent robotic systems. *Autonomous Robots*, *1*(1), 27–52. doi:10.1007/BF00735341

Breazeal, C., Kidd, C., Thomaz, A. L., Hoffman, G., & Berlin, M. (2005). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 708–713). IEEE. doi:10.1109/IROS.2005.1545011

Brenton, H., Gillies, M., Ballin, D., & Chatting, D. (2005). The uncanny valley: Does it exist? In *Proceedings of the 19th British HCI Group Annual Conference: Workshop on Human-Animated Character Interaction.* Edinburgh, UK: ACM.

Briggs, G., & Scheutz, M. (2012). Multi-modal belief updates in multi-robot human-robot dialogue interactions. In *Proceedings of the AISB/IACAP Symposium on Linguistic and Cognitive Approaches to Dialogue Agents (LaCATODA)* (pp. 67–72). Birmingham, UK: AISB.

Briggs, G., & Scheutz, M. (2014). Modeling blame to avoid positive face threats in natural language generation. In *Proceedings of the 1st Joint Session of the ACL Special Interest Groups on Natural Language Generation (SIGGEN) and Discourse and Dialogue (SIGDIAL)* (pp. 157–161). Philadelphia, PA: Association for Computational Linguistics.

DiSalvo, C. F., Gemperle, F., Forlizzi, J., & Kiesler, S. (2002). All robots are not created equal: The design and perception of humanoid robot heads. In *Proceedings of the 4th conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (pp. 321–326). doi:10.1145/778712.778756

Eyssel, F., Kuchenbrandt, D., Bobinger, S., de Ruiter, L., & Hegel, F. (2012). 'If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. In *Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 125–126). ACM. doi: 10.1145/2157689.2157717

Fong, T., Thorpe, C., & Baur, C. (2003). Collaboration, dialogue, human-robot interaction. In R. A. Jarvis & A. Zelinsky (Eds.), *Robotics Research* (pp. 255–266). Springer. doi:10.1007/3-540-36460-9_17

Foster, M. E., Gaschler, A., Giuliani, M., Isard, A., Pateraki, M., & Petrick, R. (2012). Two people walk into a bar: Dynamic multi-party social interaction with a robot agent. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction (ICMI)* (pp. 3–10). ACM.

Fraune, M., & Šabanović, S. (2014a). Negative attitudes toward minimalistic robots with intragroup communication styles. In *Proceedings of 23rd IEEE Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 1116–1121). Edinburgh, UK: IEEE.

Fraune, M., & Šabanović, S. (2014b). Robot gossip: Effects of mode of robot communication on human perceptions of robots. In *Proceedings of the 9th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 160–161). ACM. doi: 10.1145/2559636.2559832

Fukuda, T., & Sekiyama, K. (1994). Hierarchical prediction model for intelligent communication in multiple robotic systems. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 824–830). IEEE. doi:10.1109/IROS.1994.407544

Hancock, P. A., Billings, D. R., & Schaefer, K. E. (2011, September). Can you trust your robot? *Ergonomics in Design: The Quarterly of Human Factors Applications*, *19*(3), 24–29. doi: 10.1177/1064804611415045

Hayashi, K., Kanda, T., Miyashita, T., Ishiguro, H., & Hagita, N. (2008). Robot manzai: Robot conversation as a passive–social medium. *International Journal of Humanoid Robotics*, *5*(01), 67–86. doi:10.1109/ICHR.2005.1573609

Kanda, T., Ishiguro, H., Ono, T., Imai, M., & Mase, K. (2002). Multi-robot cooperation for human-robot communication. In *Proceedings of the 11th IEEE International Workshop* *on Robot and Human Interactive Communication* (pp. 271–276). IEEE. doi:10.1109/RO-MAN.2002.1045634

Kanda, T., Ishiguro, H., Ono, T., Imai, M., & Nakatsu, R. (2004, May). Effects of observation of robot-robot communication on human-robot communication. *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, *87*(5), 48–58. doi: 10.1002/ecjc.10171

Lallée, S., Pattacini, U., Lemaignan, S., Lenz, A., Melhuish, C., Natale, L., ... Dominey, P. F. (2012). Towards a platform-independent cooperative human robot interaction system: III an architecture for learning and executing actions and shared plans. *IEEE Transactions on Autonomous Mental Development*, *4*(3), 239–253. doi:10.1109/TAMD.2012.2199754

Liu, C., Ishi, C. T., Ishiguro, H., & Hagita, N. (2012). Generation of nodding, head tilting and eye gazing for human-robot dialogue interaction. In *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 285–292). ACM. doi: 10.1142/S0219843613500096

MacDorman, K. F. (2006). Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley. In *Proceedings of the 2006 ICCS/CogSci Long Symposium: Toward Social Mechanisms of Android Science* (pp. 26–29). Vancouver, BC: Cognitive Science Society.

Matsuyama, Y., Taniyama, H., Fujie, S., & Kobayashi, T. (2006). Framework of communication activation robot participating in multiparty conversation. In *Proceedings of the AAAI Fall Symposium* (pp. 68–73). ACM.

Mori, M., MacDorman (Translator), K. F., & Minato (Translator), T. (2005). The uncanny valley. *Energy*, *7*(4), 33–35.

Nadel, J., Revel, A., Andry, P., & Gaussier, P. (2004). Toward communication: First imitations in infants, low-functioning children with autism and robots. *Interaction Studies*, *5*(1), 45-74. doi:10.1075/is.5.1.04nad

Nass, C. I., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship.* MIT Press: Cambridge. doi:10.1162/coli.2006.32.3.451

Neerincx, M. A. (2007). Modelling cognitive and affective load for the design of human-machine collaboration. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics* (Vol. 4562, pp. 568–574). Springer: Berlin/Heidelberg. doi:10.1007/978-3-540-73331-7_62

Schermerhorn, P., Scheutz, M., & Crowell, C. R. (2008). Robot social presence and gender: Do females view robots differently than males? In *Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 263–270). ACM. doi: 10.1145/1349822.1349857

Scheutz, M., Schermerhorn, P., & Kramer, J. (2006). The utility of affect expression in natural language interactions in joint human-robot tasks. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction (HRI)* (pp. 226–233). ACM. doi: 10.1145/1121241.1121281

Scheutz, M., Schermerhorn, P., Kramer, J., & Anderson, D. (2007, May). First steps toward natural human-like HRI. *Autonomous Robots*, *22*(4), 411–423. doi:10.1007/s10514-006-9018-3

Tsui, K. M., Norton, A., Brooks, D. J., McCann, E., Medvedev, M. S., & Yanco, H. A. (2013). Design and development of two generations of semi-autonomous social telepresence robots. In *Proceedings of the IEEE International Conference on Technologies for Practical Robot Applications (TePRA)* (pp. 1–6). IEEE. doi:10.1109/TePRA.2013.6556360

Tsujimoto, M., Munekat, N., & Ono, T. (2013). Evaluating how the human's impression formation of robots is effected by the relation between the robots. In *Proceedings of the 1st International Conference on Human-Agent Interaction.* Tsukuba, Japan.

Walters, M. L., Syrdal, D. S., Dautenhahn, K., Te Boekhorst, R., & Koay, K. L. (2008). Avoiding the uncanny valley: Robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. *Autonomous Robots*, *24*(2), 159–178. doi: 10.1007/s10514-007-9058-3

Wang, J. W. J. (1994). On sign-board based inter-robot communication in distributed robotic systems. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation (ICRA)* (p. 1045-1050). IEEE. doi:10.1109/ROBOT.1994.351219

Williams, T., Briggs, G., Oosterveld, B., & Scheutz, M. (2015). Going beyond command-based instructions: Extending robotic natural language interaction capabilities. In *Proceedings of 29th AAAI Conference on Artificial Intelligence* (p. 1387-1393). Austin, TX: ACM.

Williams, T., Briggs, P., Pelz, N., & Scheutz, M. (2014). Is robot telepathy acceptable? Investigating effects of nonverbal robot-robot communication on human-robot interaction. In *Proceedings of the 23rd IEEE Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 886–891). IEEE. doi:10.1109/ROMAN.2014.6926365

Authors' names and contact information: Tom Williams, Priscilla Briggs, and Matthias Scheutz are with the Human-Robot Interaction Laboratory at Tufts University in Medford, MA, USA. {williams, pbriggs, mscheutz}@cs.tufts.edu