4th International Symposium on New Frontiers in Human-Robot Interaction
21–22 April 2015, Canterbury, UK
M. Salem, A. Weiss, P. Baxter, K. Dautenhahn (Eds.)

Robots Have Needs Too: People Adapt Their Proxemic Preferences to Improve Autonomous Robot Recognition of Human Social Signals

Ross Mead[1] and **Maja J Matarić**[2]

Abstract. An objective of autonomous socially assistive robots is to meet the needs and preferences of human users. However, this can sometimes be at the expense of the robot's own ability to understand *social signals* produced by the user. In particular, human preferences of distance (*proxemics*) to the robot can have significant impact on the performance rates of its automated speech and gesture recognition systems. In this work, we investigated how user proxemic preferences changed to improve the robot's understanding human social signals. We performed an experiment in which a robot's ability to understand social signals was artificially varied, either *uniformly* or *attenuated* across distance. Participants ($N = 100$) instructed a robot using speech and pointing gestures, and provided their proxemic preferences before and after the interaction. We report two major findings: 1) people predictably underestimate (based on a Power Law) the distance to the location of robot peak performance; and 2) people adjust their proxemic preferences to be near the *perceived* location of robot peak performance. This work offers insights into the dynamic nature of human-robot proxemics, and has significant implications for the design of social robots and robust autonomous robot proxemic control systems.

1 Introduction

A social robot utilizes natural communication mechanisms, such as speech and gesture, to autonomously interact with humans to accomplish some individual or joint task [2]. The growing field of socially assistive robotics (SAR) is at the intersection of social robotics and assistive robotics that focuses on non-contact human-robot interaction (HRI) aimed at monitoring, coaching, teaching, training, and rehabilitation domains [4]. Notable areas of SAR include robotics for older adults, children with autism spectrum disorders, and people in post-stroke rehabilitation, among others [25, 17].

Consequently, SAR constitutes an important subfield of robotics with significant potential to improve health and quality of life. Because the majority of SAR contexts investigated to date involve one-on-one face-to-face interaction between the robot and the user, how the robot understands and responds to the user is crucial to successful autonomous social robots [1], in SAR contexts and beyond.

One of the most fundamental social behaviors is *proxemics*, the social use of space in face-to-face social encounters [5]. A mobile social robot must position itself appropriately when interacting with the user. However, robot position has a significant impact on the robot's *performance*—in this work, performance is measured by automated speech and gesture recognition rates. Just like electrical signals, human *social signals* (e.g., speech and gesture) are *attenuated* (lose signal strength) based on distance, which dramatically changes the way in which automated recognition systems detect and identify the signal; thus, a proxemic control system that often varies its location and, thus, creates signal attenuation, can be a defining factor in the success or failure of a social robot [16].

In our previous work [16] (described in detail in Section 2.2), we modeled social robot performance attenuated by distance, which was then used to implement an autonomous robot proxemic controller that maximizes its performance during face-to-face HRI; however, this work begged the question as to whether or not people would accept a social robot that positions itself in a way that differs from traditional user proxemic preferences. Would users naturally change their proxemic preferences if they observed differences in robot performance in different proxemic configurations, or would their proxemic preferences persist, mandating that robot developers must improve autonomous speech and gesture recognition systems before social and socially assistive robot technology can be deployed in the real world? This question is the focus of the investigation reported here.

[1] University of Southern California, USA, email: rossmead@usc.edu
[2] University of Southern California, USA, email: mataric@usc.edu

2 Background

The anthropologist Edward T. Hall [5] coined the term "proxemics", and, in [6], proposed that proxemics lends itself well to being analyzed with performance (as measured through sensory experience) in mind. Proxemics has been studied in a variety of ways in HRI; here, we constrain our review of related work to that of *autonomous* HRI[3].

2.1 Comfort-based Proxemics in HRI

The majority of proxemics work in HRI focuses on maximizing user *comfort* during a face-to-face interaction. The results of many human-robot proxemics studies have been consolidated and normalized in [28], reporting mean distances of 0.49–0.71 meters using a variety of robots and conditions. Comfort-based proxemic preferences between humans and the PR2 robot[4] were investigated in [24], reporting mean distances of 0.25–0.52 meters; in [16], we investigated the same preferences using the PR2 in a conversational context, reporting a mean distance of 0.94 meters. Farther proxemic preferences have been measured in [18] and [26], reporting mean distances of 1.0–1.1 meters and 1.7–1.8 meters, respectively.

[3] There is a myriad of related work reporting how humans adapt to various technologies, but this is beyond the scope of this work. For a review, see [8].
[4] https://www.willowgarage.com/pages/pr2/overview

However, results in our previous work [16] suggest that autonomous speech and gesture recognition systems do not perform well using comfort-based proxemic configurations. Speech recognition performed adequately at distances less than 2.5 meters, and face and hand gesture recognition performed adequately at distances of 1.5–2.5 meters; thus, given current technologies, distances for mutual recognition of these social signals is between 1.5 and 2.5 meters, at and beyond the far end of comfort-based proxemic preferences.

2.2 Performance-based Proxemics in HRI

Our previous work utilized advancements in markerless motion capture (specifically, the Microsoft Kinect) to automatically extract proxemic features based on metrics from the social sciences [11, 14]. These features were then used to recognize spatiotemporal interaction behaviors, such as the initiation, acceptance, aversion, and termination of an interaction [12, 14]. These investigations offered insights into the development of proxemic controllers for autonomous social robots, and suggested an alternative approach to the representation of proxemic behavior that goes beyond simple distance and orientation [13]. A probabilistic framework for autonomous proxemic control was proposed in [15, 10] that considers *performance* by maximizing the sensory experience of each agent (human or robot) in a co-present social encounter. The methodology established an elegant connection between previous approaches and illuminated the functional aspects of proxemic behavior in HRI [13], specifically, the impact of spacing on speech and gesture behavior recognition and production. In [16], we formally modeled (using a dynamic Bayesian network [9]) autonomous speech and gesture recognition systems as a function of distance and orientation between a social robot and a human user, and implemented the model as an autonomous proxemic controller, which was shown to maximize robot performance in HRI.

However, while our approach to proxemic control *objectively* maximized the performance of the robot, it also resulted in proxemic configurations that are atypical for human-robot interactions (e.g., positioning itself farther or nearer to the user than preferred). Thus, the question arose as to whether or not people would *subjectively* adopt a technology that places performance over preference, as it might place a burden on people to change their own behaviors to make the technology function adequately.

2.3 Challenges in Human Spatial Adaptation

For humans to adapt their proxemic preferences to a robot, they must be able to accurately identify regions in which the robot is performing well; however, errors in human distance estimation increase nonlinearly with increases in distance, time, and uncertainty [19]. Fortunately, the relationship between human distance estimation and each of these factors is very well represented by Steven's Power Law, ax^b, where x is distance [19, 23]. Unfortunately, these relationships are reported for distances of 3–23 meters, which are farther away than in those with which we are concerned for face-to-face HRI—thus, we cannot use the reported model parameters and must derive our own.

In this work, we investigate how user proxemic preferences change in the presence of a social robot that is recognizing and responding to instructions provided by a human user. Robot performance (ability to understand speech and gesture) is artificially attenuated to expose participants to success and failure scenarios while interacting with the robot. In Section 3, we describe the overall setup in which our investigation took place. In Section 4, we outline the specific procedures, conditions, hypotheses, and participants of our experiment.

3 Experimental Setup

3.1 Materials

The experimental robotic system used in this work was the Bandit upper-body humanoid robot[5] [Figure 1]. Bandit has 19 degrees of freedom: 7 in each arm (shoulder forward-and-backward, shoulder in-and-out, elbow tilt, elbow twist, wrist twist, wrist tilt, grabber open-and-close; left and right arms), 2 in the head (pan and tilt), 2 in the lips (upper and lower), and 1 in the eyebrows. These degrees of freedom allow Bandit to be expressive using individual and combined motions of the head, face, and arms. Mounted atop a Pioneer 3-AT mobile base[6], the entire robot system is 1.3 meters tall.

A Bluetooth PlayStation 3 (PS3) controller served as a remote control interface with the robot. The controller was used by the experimenter (seated behind a one-way mirror [Figure 2]) to step the robot through each part of the experimental procedure (described in Section 4.1)—the decisions and actions taken by the robot during the experiment were completely autonomous, but the timing of its actions were controlled by the press of a "next" button. The controller was also used to record distance measurements during the experiment, and to provide ground-truth information to the robot as to what the participant was communicating (however, the robot autonomously determined how to respond based on the experimental conditions described in Section 4.2).

Four small boxes were placed in the room, located at 0.75 meters and 1.5 meters from the centerline on each side (left and right) of the participant [Figure 2]. During the experiment (described in Section 4.1), the participant instructed the robot to look at these boxes. Each box was labeled with a unique shape and color; in this experiment, the shapes and colors matched the buttons on the PS3 controller: a green triangle, a red circle, a blue cross, and a purple square. This allowed the experimenter to easily indicate to the robot to which box the user was attending (i.e., "ground-truth").

A laser rangefinder on-board the robot was used to measure the distance from the robot to the participant's legs at all times.

Figure 1. The Bandit upper-body humanoid robot.

[5] http://robotics.usc.edu/interaction/?l=Laboratory:Robots#BanditII
[6] http://www.mobilerobots.com/ResearchRobots/P3AT.aspx

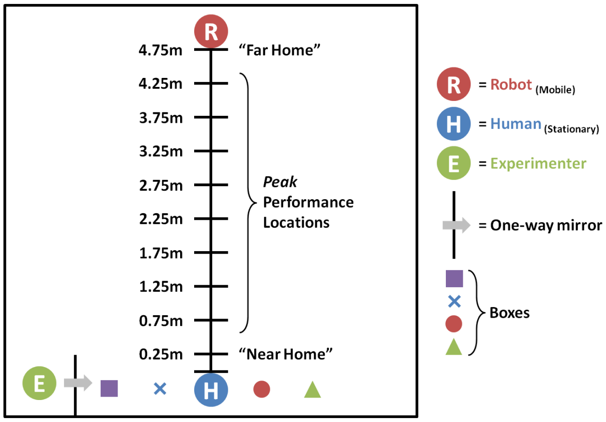

Figure 2. The experimental setup.

3.2 Robot Behaviors

The robot autonomously executed three primary behaviors throughout the experiment: 1) forward and backward base movement, 2) maintaining eye contact with the participant, and 3) responding to participant instructions with head movements and audio cues.

Robot base movement was along a straight-line path directly in front of the participant, and was limited to distances of 0.25 meters (referred to as the "near home" location) and 4.75 meters (referred to as the "far home" location); it returned repeatedly to these "home" locations throughout the experiment. Robot velocity was proportional to the distance to the goal location; the maximum robot speed was 0.3 m/s, which people find acceptable [22].

As the robot moved, it maintained eye contact with the participant. The robot has eyes, but they are not actuated, so the robot's head pitched up or down depending on the location of the participant's head, which was determined by the distance to the participant (from the on-board laser) and the participant's self-reported height. We note that prolonged eye contact from the robot often results in user preferences of increased distance in HRI [24, 18].

The robot provided head movement and audio cues to indicate whether or not it understood instructions provided by the participant (described in Section 4.1.2). If the robot understood the instructions, it provided an *affirmative response* (looking at a box); if the robot did not understand the instructions, it provided a *negative response* (shaking its head). With each head movement, one of two affective sounds were also played to supplement the robot's response; affective sounds were used because robot speech influences proxemic preferences and would have introduced a confound in the experiment [29].

4 Experimental Design

With the described experimental setup, we performed an experiment to investigate user perceptions of robot performance attenuated by distance and its effect on proxemic preferences.

4.1 Experimental Procedure

Participants (described in Section 4.4) were greeted at the door entering the private experimental space, and were informed of and agreed to the nature of the experiment and their rights as a participant, which included a statement that the experiment could be halted at any time.

Participants were then instructed to stand with their toes touching a line on the floor, and were asked to remain there for the duration of the experiment [Figure 2]. The experimenter then provided instructions about the task the participant would be performing.

Participants were introduced to the robot, and were informed that all of its actions were completely autonomous. Participants were told that the robot would be moving along a straight line throughout the duration of the experiment; a brief demonstration of robot motion was provided, in which the robot autonomously moved back and forth between distances of 3.0 meters and 4.5 meters from the participant, allowing them to familiarize themselves with the robot motion. Participants were told that they would be asked about some of their preferences regarding the robot's location throughout the experiment.

Participants were then informed that they would be instructing the robot to look at any one of four boxes (of their choosing) located in the room [Figure 2], and that they could use speech (in English) and pointing gestures. A vocabulary for robot instructions was provided: for speech, participants were told they could say the words "look at" followed by the name of the shape or color of each box (e.g., "triangle", "circle", "blue", "purple", etc.); for pointing gestures, participants were asked to use their left arm to point to boxes located on their left, and their right arm to point to boxes on their right. This vocabulary was provided to minimize any perceptions the person might have that the robot simply did not understand the words or gestures that they used; thus, the use of the vocabulary attempted to maximize the perception that any failures of the robot were due to other factors.

Participants were told that they would repeat this instruction procedure to the robot many times, and that the robot would indicate whether or not it understood their instructions each time using the head movements and audio cues described in Section 3.2.

Participants had an opportunity to ask the experimenter any clarifying questions. Once participant understanding was verified, we proceeded with the experiment.

4.1.1 Pre-interaction Proxemic Measures (pre)[7]

The robot autonomously moved to the "far home" location [Figure 2]. Participants were told that the robot would be approaching them, and to say out loud the word "stop" when the robot reached the ideal location at which the participant would have a *face-to-face conversation*[8] with the robot. This pre-interaction proxemic preference from the "far home" location is denoted as pre_{far}.

When the participant was ready, the experimenter pressed a PS3 button to start the robot moving. When the participant said "stop", the experimenter pressed another button to halt robot movement. The experimenter pressed another button to record the distance between the robot and the participant, as measured by the on-board laser.

Once the pre_{far} distance was recorded, the experimenter pressed another button, and the robot autonomously moved to the "near home" location [Figure 2]; the participant was informed that the robot would be approaching to this location and would stop on its own. The process was repeated with the robot backing away from the participant, and the participant saying "stop" when it reached the ideal location for conversation. This pre-interaction proxemic preference from the "near home" location is denoted as pre_{near}.

[7] Measures are provided inline with the experimental procedure to provide an order of events as they occurred in the experiment.

[8] Related work in human-robot proxemics asks the participant about locations at which they feel *comfortable* [24], yielding proxemic preferences very near to the participant. Our general interest is in face-to-face human-robot conversational interaction, with proxemic preference farther from the participant [16, 26, 27], hence the choice of wording.

From pre_{far} and pre_{near}, we calculated and recorded the average pre-interaction proxemic preference, denoted as pre[9].

4.1.2 Interaction Scenario

After determining pre-interaction proxemic preferences, the robot returned to the "far home" location. The experimenter then repeated to participants the instructions about the task they would be performing with the robot. When participants verified that they understood the task and indicated that they were ready, the experimenter pressed a button to proceed with the task.

The robot autonomously visited ten pre-determined locations [Figure 2]. At each location, the robot responded to instructions from the participant to look at one of four boxes located in the room [Figure 2]. Five instruction-response interactions were performed at each location, after which the robot moved to the next location along its path; thus, each participant experienced a total of 50 instruction-responses interactions. Robot goal locations were in 0.5-meter intervals inclusively between the "near home" location (0.25 meters) and "far home" location (4.75 meters) along a straight-line path in front of the participant [Figure 2]. Locations were visited in a sequential order; for half of the participants, the robot approached from the "far home" location (i.e., farthest-to-nearest order), and, for the other half of participants, the robot backed away from "near home" location (i.e., nearest-to-farthest order); this was done to reduce any ordering effects [19].

To controllably simulate social signal attenuation at each location, robot performance was artificially manipulated as a function of the distance to the participant (described in Section 4.2). After each instruction provided by the participant, the experimenter provided to the robot (via a remote control interface) the ground-truth of the instruction; the robot then determined whether or not it would have understood the instruction based on a prediction from a performance vs. distance curve (specified by the assigned experimental condition described in Section 4.2), and provided either an *affirmative response* or a *negative response* to the participant indicating its successful or failed understanding of the instruction, respectively.

The entire interaction scenario lasted 10-15 minutes.

4.1.3 Post-interaction Proxemic Measures ($post$)

After the robot visited each of the ten locations, it autonomously returned to the "far home" location. The experimenter then repeated the procedure for determining proxemic preferences described in Section 4.1.1. This process generated post-performance proxemic preferences from the "far home" and "near home" locations, as well as their average, denoted $post_{far}$, $post_{near}$, and $post$[10], respectively.

4.1.4 Perceived Peak Location Measures ($perc$)

Finally, after collecting post-interaction proxemic preferences, the experimenter repeated the procedure described in Section 4.1.1 to determine participant perceptions of the location of peak performance. This process generated perceived peak performance locations from the "far home" and "near home" locations, as well as their average, denoted $perc_{far}$, $perc_{near}$, and $perc$[11], respectively.

[9]Post-hoc analysis revealed no statistically significant difference between pre_{far} and pre_{near} measurements, hence why we rely on pre.

[10]Post-hoc analysis revealed no statistically significant difference between $post_{far}$ and $post_{near}$ measurements, hence why we rely on $post$.

[11]Post-hoc analysis revealed no statistically significant difference between $perc_{far}$ and $perc_{near}$ measurements, hence why we rely on $perc$.

4.2 Experimental Conditions

We considered two performance vs. distance conditions; 1) a **"uniform performance" condition**, and 2) an **"attenuated performance" condition**. Overall robot performance for each condition was held at a constant 40%[12]—that is, for each participant, the robot provided 20 affirmative responses and 30 negative responses distributed across 50 instructions. The way in which these responses were distributed across locations varied between conditions.

In the **uniform performance condition**, robot performance was the same (40%) across across all locations [Figures 3 and 4]. Thus, at each of the ten locations visited, the robot provided two affirmative and three negative responses, respectively. This condition served as a baseline of participant proxemic preferences within the task.

In the **attenuated performance condition**, robot performance varied with distance proportional to a Gaussian distribution centered a location of "peak performance" ($M = peak$, $SD = 1.0$) [Figures 3 and 4]. Due to differences in pre-interaction proxemic preferences, we could not select a single value for $peak$ that provided a similar experience between participants without introducing other confounding factors (e.g., the $peak$ not being at a location that the robot visited or distances beyond the "home" locations). To alleviate this, we opted to select multiple peak performance locations, exploring the space of human responses to robot performance differences at a variety of distances. We selected the eight locations non-inclusively between the "near home" and "far home" locations as the peak performance locations [Figure 2]; the "near home" and "far home" locations were not included in the set of peaks to ensure that participants were always exposed to an actual *peak* in performance, rather than just a *trend*. Peak performance locations were varied between participants.

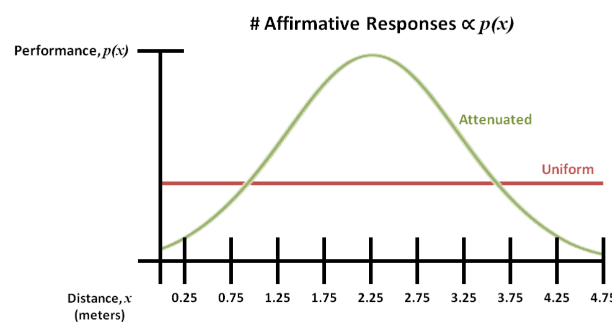

Figure 3. The performance curves of the **uniform** and **attenuated** conditions. In this example, $peak = 2.25$ (in meters), so the attenuated performance curve parameters is $M = peak = 2.25$, $SD = 1.0$. The number of affirmative responses at a distance, x, from the user is proportional to $p(x)$, the evaluation of the performance curve at x.

The distribution of affirmative responses for all conditions is presented in Figure 4. The number of affirmative responses was normalized to 20 (40%) to ensure a consistent user experience of overall robot performance across all conditions. In the **attenuated performance condition**, the number of affirmative responses at $peak$ was always the 5 (i.e., perfect performance), and the number of affirmative responses at other locations were always less than that of the peak to ensure that participants were exposed to an actual peak. At each location, the order in which the five responses were provided was random.

[12]This value was selected because it is an average performance rate predicted by our results in [16] for typical human-robot proxemic preferences.

Performance Condition		# *Affirmative Responses vs. Distance*									
Distance, x (meters) →		0.25	0.75	1.25	1.75	2.25	2.75	3.25	3.75	4.25	4.75
Uniform →		2	2	2	2	2	2	2	2	2	2
Attenuated $peak$ → (meters)	***0.75***	4	***5***	4	3	1	1	1	1	0	0
	1.25	3	4	***5***	4	3	1	0	0	0	0
	1.75	1	3	4	***5***	4	3	0	0	0	0
	2.25	0	1	3	4	***5***	4	3	0	0	0
	2.75	0	0	0	3	4	***5***	4	3	1	0
	3.25	0	0	0	0	3	4	***5***	4	3	1
	3.75	0	0	0	0	1	3	4	***5***	4	3
	4.25	0	0	1	1	1	1	3	4	***5***	4

Figure 4. The distribution of affirmative responses provided by the robot across conditions. Manipulated values are highlighted in **bold italics**.

4.3 Experimental Hypotheses

Within these conditions, we had three central hypotheses:

H1: In the **uniform performance condition**, there will be no significant change in participant proxemic preferences.

H2: In the **attenuated performance conditions**, participants will be able to identify a relationship between robot performance and human-robot proxemics.

H3: In the **attenuated performance conditions**, participants will adapt their proxemic preferences to improve robot performance.

4.4 Participants

We recruited 100 participants (50 male, 50 female) from our university campus community. Participant race was diverse (67 white/Caucasian, 26 Asian, 3 black/African-American, 3 Latino/Latina, and 1 mixed-race). All participants reported proficiency in English and had lived in the United States for at least two years (i.e., acclimated to U.S. culture). Average age (in years) of participants was 22.26 ($SD = 4.31$), ranging from 18 to 39. Based on a seven-point scale, participants reported moderate familiarity with technology ($M = 3.98, SD = 0.85$). Average participant height (in meters) was 1.74 ($SD = 0.10$), ranging from 1.52 to 1.93. Related work reports how human-robot proxemics is influenced by gender and technology familiarity [24], culture [3], and height [7, 21].

The 100 participants were randomly assigned to a performance condition, with $N = 20$ in the **uniform performance condition** and $N = 80$ in the **attenuated performance condition**. In the **attenuated performance condition**, the 80 participants were randomly assigned one of the eight peak performance locations (described in Section 4.2) with $N = 10$ for each $peak$. Neither the participant nor the experimenter was aware of the condition assigned.

5 Results and Discussion

We analyzed data collected in our experiment to test our three hypotheses (described in Section 4.3), and evaluated their implications for autonomous social robots and human-robot proxemics.

To provide a baseline of our robot for comparison in general human-robot proxemics, we consolidated and analyzed pre-interaction proxemic preferences (pre) across all conditions ($N = 100$), as the data had not been influenced by robot performance. The participant pre-interaction proxemic preference (in meters) was determined to be 1.14 ($SD = 0.49$) for our robot system, which is consistent with [18] and our previous work [16], but twice as far away as related work has reported for robots of a similar form factor [28, 24].

5.1 H1: Pre- vs. Post-interaction Locations

To test **H1**, we compared average pre-interaction proxemic preferences (pre) to average post-interaction proxemic preferences ($post$) of participants in the **uniform performance condition**.

A paired t-test revealed a statistically significant change in participant proxemic preferences between pre ($M = 1.12, SD = 0.51$) and $post$ ($M = 1.39, SD = 0.63$); $t(38) = 1.49, p = 0.02$. Thus, our hypothesis **H1** is rejected.

The rejection of this hypothesis does not imply a failure of the experimental procedure, but, rather, provides important insights that must be considered for subsequent analyses (and for related work in proxemics). This result suggests that there might be something about the context of the interaction scenario itself that influenced participant proxemic preferences. To address any influence the interaction scenario might have on subsequent analyses, we define a *contextual offset*, θ, as the average difference in participant post-interaction and pre-interaction proxemic preferences ($M = 0.27, SD = 0.48$); this θ value will be subtracted from ($post - pre$) values in Section 5.3 to normalize for the interaction context.

5.2 H2: Perceived vs. Actual Peak Locations

To test **H2**, we compared participant perceived locations of peak performance ($perc$) to actual locations of peak performance ($peak$) in the **attenuated performance conditions** [Figure 5].

Steven's Power Law, ax^b, has previously been used to model human distance estimation as a function of actual distance [19], and is generally well representative of human-perceived vs. actual stimuli [23]. However, existing Power Laws relevant to our work only seem to pertain to distances of 3–23 meters, which are beyond the range of the natural face-to-face communication with which we are concerned. Thus, our goal here is to model our own experimental data to establish a Power Law for $perc$ vs. $peak$ at locations more relevant to HRI (0.75–4.25 meters), which we can then evaluate to test **H2**.

Immediate observations of our data suggested that the data appear to be heteroscedastic [Figure 5]—in this case, the variance seems to increase with distance from the participant, which means we should not use traditional statistical tests. The Breusch-Pagan test for non-constant variance (NCV) confirmed this intuition; $\chi^2(1, N = 100) = 15.79, p < 0.001$. A commonly used and accepted approach to alleviate our heteroscedasticity is to transform the $perc$ and $peak$ data to a log-log scale. While not applicable to all datasets, this approach served as an adequate approximation for our purposes [Figure 6]; it also enabled us to perform a regression analysis to determine parameter values for the Power Law coefficient and exponent, $a = 1.3224$ and $b = 0.5132$, respectively. With these parameters, we identified that $peak$ was a strongly correlated and very significant predictor of $perc$; $R^2 = 0.4951, F(1, 78) = 76.48, p < 0.001$. Thus, our hypothesis **H2** is supported.

This result suggests that people are able to identify a relationship between robot performance and human-robot proxemics, but they will predictably underestimate the distance, x, to the location of peak performance based on the Power Law equation $1.3224x^{0.5132}$. While human estimation of the location of peak performance is suboptimal, it is possible that repeated exposure to the robot over multiple sessions might yield more accurate results. This follow-up hypothesis will be formally tested in a planned longitudinal study in future work (described in Section 6).

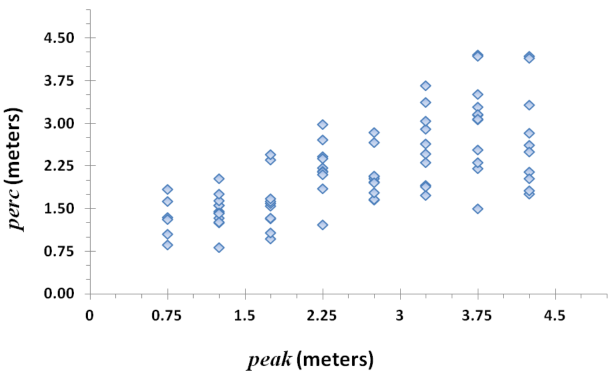

Figure 5. Participant perceived location of robot peak performance ($perc$) vs. actual location of robot peak performance ($peak$). Note the heteroscedasticity of the data, which prevents us from performing traditional statistical analyses without first transforming the data (shown in Figure 6).

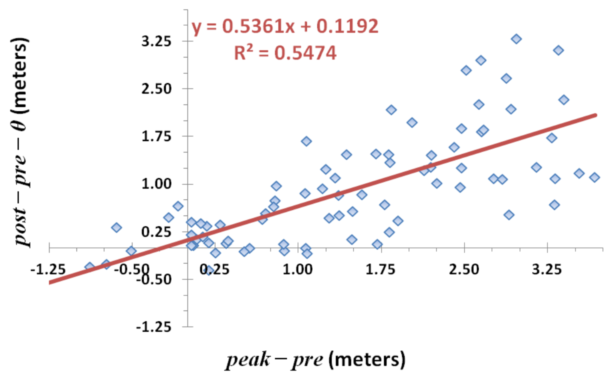

Figure 7. Changes in participant pre-/post-interaction proxemic preferences (pre and $post$, respectively; θ is the *contextual offset* defined in Section 5.1) vs. distance from participant pre-interaction proxemic preference (pre) to the actual location of robot peak performance ($peak$).

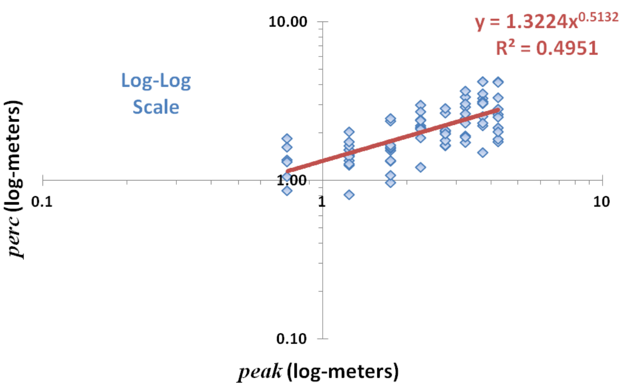

Figure 6. Participant perceived location of robot peak performance ($perc$) vs. actual location of robot peak performance ($peak$) on a log-log scale, reducing the effects of heteroscedasticity and allowing us to perform regression to determine parameters of the Power Law, ax^b.

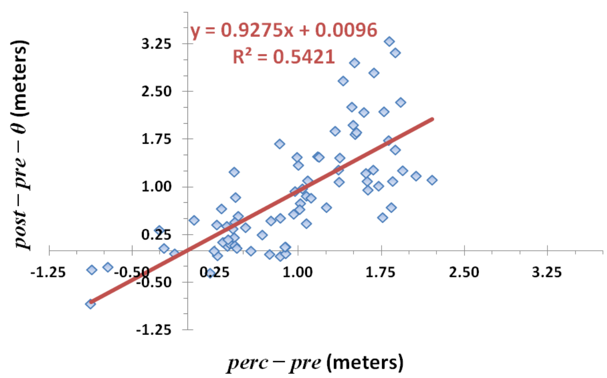

Figure 8. Changes in participant pre-/post-interaction proxemic preferences (pre and $post$, respectively; θ is the *contextual offset* defined in Section 5.1) vs. distance from participant pre-interaction proxemic preference (pre) to the perceived location of robot peak performance ($perc$).

5.3 H3: Preferences vs. Peak Locations

To test **H3**, we compared changes in participant pre-/post-interaction proxemic preferences ($post - pre - \theta$) to the distance from the participant pre-interaction proxemic preference to either a) the actual location of robot peak performance ($peak - pre$) [Figure 7], or b) the perceived location of robot peak performance ($perc - pre$) [Figure 8], both in the **attenuated performance conditions**.

Data for ($post - pre - \theta$) vs. both ($peak - pre$) and ($perc - pre$) were heteroscedastic, as indicated by Breusch-Pagan NCV tests: $\chi^2(1, N = 100) = 18.81, p < 0.001$; and $\chi^2(1, N = 100) = 13.55, p < 0.001$; respectively. This is intuitive, as the data for perceived ($perc$) vs. actual ($peak$) locations of peak performance were also heteroscedastic [Figure 5]. The log-transformation approach that we used in Section 5.2 did not perform well in modeling these data; thus, we needed to use an alternative approach. We opted to utilize a Generalized Linear Model [20] because it allowed us to model the variance of each measurement separately as a function of predicted values and, thus, perform appropriate statistical tests for significance.

We first modeled changes in participant proxemic preferences ($post - pre - \theta$) vs. distance from pre-interaction proxemic preference to the actual location of peak performance ($peak - pre$). In the ideal situation (for the robot), these match one-to-one—in other words, the participant meets the needs of the robot entirely by changing proxemic preferences to be centered at the peak of robot performance. Unfortunately for the robot, this was not the case. We detected a strongly correlated and statistically significant relationship between participant proxemic preference change and distance from pre-interaction preference to the peak location ($R^2 = 0.5474, \beta = 0.5361, t(98) = 9.71, p < 0.001$), but participant preference change only got the robot approximately halfway ($\beta = 0.5361$) to its location of peak performance [Figure 7]. Why is this?

Recall that results reported in Section 5.2 suggested that, while people do perceive a relationship between robot performance and distance, their ability to accurately identify the location of robot peak performance diminishes based on the distance to it as governed by a Power Law. Were participants *trying* to maximize robot performance, but simply adapting their preferences to a suboptimal location?

We investigated this question by considering changes in participant proxemic preferences ($post - pre - \theta$) vs. distance from pre-interaction proxemic preference to the perceived location of peak performance ($perc - pre$). If the participant was adapting their proxemic preferences to accommodate the needs of the robot, then these should match one-to-one. A Generalized Linear Model was fit to these data, and yielded a strongly correlated and statistically significant relationship between changes in proxemic preferences and perceptions of robot performance ($R^2 = 0.5421, \beta = 0.9275, t(98) = 9.61, p < 0.001$) [Figure 8]. Thus, our hypothesis **H3** is supported.

The near one-to-one relationship ($\beta = 0.9275$) between post-interaction proxemic preferences and participant perceptions of robot peak performance is compelling, suggesting that participants adapted their proxemic preferences almost entirely to improve robot performance in the interaction.

5.4 Discussion

These results have significant implications for the design of social robots and autonomous robot proxemic control systems, specifically, in that people's proxemic preferences will likely change as the user interacts with and comes to understand the needs of the robot.

As illustrated in our previous work [16], the locations of on-board sensors for social signal recognition (e.g., microphones and cameras), as well as the automated speech and gesture recognition software used, can have significant impacts on the performance of the robot in autonomous face-to-face social interactions. As our now-reported results suggest that people will adapt their behavior in an effort to improve robot performance, it is anticipated that human-robot proxemics will vary between robot platforms with different hardware and software configurations based on factors that are 1) not specific to the user (unlike culture [3], or gender, personality, or familiarity with technology [24]), 2) not observable to the user (unlike height [7, 21], amount of eye contact [24, 18], or vocal parameters [29]), or 3) not observable to the robot developer. User understanding of the relationship between robot performance and human-robot proxemics is a latent factor that only develops through repeated interactions with the robot (perhaps expedited by the robot communicating its predicted error); fortunately, our results indicate that user understanding will develop in a predictable way. Thus, it is recommended that social robot developers consider and perhaps model robot performance as a function of conditions that might occur in dynamic proxemic interactions with human users to better predict and accommodate how the people will actually use the technology. This dynamic relationship, in turn, will enable more rich autonomy for social robots by improving the performance of their own automated recognition systems.

If developers adopt models of robot performance as a factor contributing to human-robot proxemics, then it follows that proxemic control systems might be designed to expedite the process of autonomously positioning the robot at an optimal distance from the user to maximize robot performance while still accommodating the initial personal space preferences of the user. This was the focus of our previous work [16], which treated proxemics as an optimization problem that considers the production and perception of social signals (speech and gesture) as a function of distance and orientation. Recall that an objective of the now-reported work was to address questions regarding whether or not users would accept a robot that positions itself in locations that might differ from their initial proxemic preferences. The results in this work (specifically, in Section 5.3) support the notion that user proxemic preferences will change through interactions with the robot as its performance is observed, and that the new user proxemic preference will be at the *perceived* location of robot peak performance. An extension of this result is that, through repeated interactions, user proxemic preferences will further adapt and eventually converge to the *actual* location of robot peak performance, a hypothesis that we will investigate in future work.

6 Future Work

Our experimental conditions (described in Section 4.2) were specifically selected to strongly expose a relationship (if one existed) between human proxemic preferences and robot performance—the robot achieved perfect success rates (100%) at "peak" locations and perfect failure rates (0%) at other locations, and these success/failure rates were distributed proportional to a Gaussian distribution with constant variance. Now that we have identified that a relationship exists, our next steps will examine how the relationship changes over time or with other related factors. A longitudinal study over multiple sessions will be conducted to determine if changes in preferences persist from one interaction to the next, and if user proxemic preferences will continue to adapt and eventually converge to locations of robot peak performance through repeated interactions. Other future work will follow the same experimental procedure described in Section 4.1, but will adjust the **attenuated performance condition** (described in Section 4.2) to consider how the relationship changes with 1) distributions of lower or higher variance, 2) lower maximum performance or higher minimum performance, 3) more realistic non-Gaussian distributions, and 4) the interactions between distributions of actual multimodal recognition systems [16].

This perspective opens up a whole new theoretical design space of human-robot proxemic behavior. The general question is, "How will people adapt their proxemic preferences in any given *performance field*?", in which performance varies with a variety of factors, such as distance, orientation, and environmental interference. The follow-up question then asks, "How can the robot expedite the process of establishing an appropriate human-robot proxemic configuration within the performance field without causing user discomfort?" This will be a focus of future work, and will extend our prior work on modeling human-robot proxemics to improve robot proxemic controllers [16].

7 Summary and Conclusions

An objective of autonomous socially assistive robots is to meet the needs and preferences of a human user [4]. However, this can sometimes be at the expense of the robot's own ability to understand social signals produced by the user. In particular, human proxemic preferences with respect to a robot can have significant impacts on the performance rates of its automated speech and gesture recognition systems [16]. This means that, for a successful interaction, the robot has needs too—and these needs might not be consistent with and might require changes in the proxemic preferences of the human user.

In this work, we investigated how user proxemic preferences changed to improve the robot's understanding of human social signals (described in Section 4). We performed an experiment in which a robot's performance was artificially varied, either *uniformly* or *attenuated* across distance. Participants ($N = 100$) instructed a robot using speech and pointing gestures, and provided their proxemic preferences before and after the interaction.

We report two major findings. First, people predictably underestimate the distance to the location of robot peak performance; the relationship between participant perceived and actual distance to the location of peak performance is represented well by a Power Law (described in Section 5.2). Second, people adjust their proxemic preferences to be near the *perceived* location of maximum robot understanding (described in Section 5.3). This work offers insights into the dynamic nature of human-robot proxemics, and has significant implications for the design of social robots and robust autonomous robot proxemic control systems (described in Section 5.4).

Traditionally, we focus on our attention on ensuring the robot is meeting the needs of the user with little regard to the impact it might have on the robot itself; it is often an afterthought, or something that we, as robot developers, have to "fix" with our systems. While robot developers will continue to improve upon our autonomous systems, our results suggest that even novice users are willing to adapt their behaviors in an effort to help the robot better understand and perform its tasks. Automated recognition systems are not and will likely never be perfect, but this is no reason to delay the development, deployment, and benefits of social and socially assistive robot technologies. Robots have needs too, and human users will attempt to meet them.

ACKNOWLEDGEMENTS

This work is supported in part by an NSF Graduate Research Fellowship, the NSF National Robotics Initiative (IIS-1208500), and an NSF CNS-0709296 grant.

We thank Aditya Bhatter, Lizhi Fan, Jonathan Frey, Akash Metawala, Kedar Prabhu, and Cherrie Wang for their assistance in recruiting participants and conducting the experiment.

REFERENCES

[1] C. Breazeal, 'Social interactions in hri: The robot view', *IEEE Transactions on Man, Cybernetics and Systems*, **34**(2), 181–186, (2003).

[2] C. Breazeal, *Designing Sociable Robots*, MIT Press, Cambridge, Massachusetts, 2004.

[3] G. Eresha, M. Haring, B. Endrass, E. Andre, and M. Obaid, 'Investigating the influence of culture on proxemic behaviors for humanoid robots', in *22nd IEEE International Symposium on Robot and Human Interactive Communication*, RO-MAN 2013, pp. 430–435, (2013).

[4] D.J. Feil-Seifer and M.J. Matarić, 'Defining socially assistive robotics', in *International Conference on Rehabilitation Robotics*, ICRR'05, pp. 465–468, Chicago, Illinois, (2005).

[5] E. T. Hall, *The Silent Language*, Doubleday Company, New York, New York, 1959.

[6] E.T Hall, 'A system for notation of proxemic behavior', *American Anthropologist*, **65**, 1003–1026, (1963).

[7] Y. Hiroi and A. Ito, *Influence of the Size Factor of a Mobile Robot Moving Toward a Human on Subjective Acceptable Distance*.

[8] P.H. Kahn, *Technological Nature: Adaptation and the Future of Human Life*, MIT Press, Cambridge, Massachusetts, 2011.

[9] D. Koller and N. Friedman, *Probabilistic Graphical Models*, MIT Press, Cambridge, Massachusetts, 2009.

[10] R. Mead, 'Space, speech, and gesture in human-robot interaction', in *Doctoral Consortium of International Conference on Multimodal Interaction*, ICMI'12, pp. 333–336, Santa Monica, California, (2012).

[11] R. Mead, A. Atrash, and M. J. Matarić, 'Proxemic feature recognition for interactive robots: Automating metrics from the social sciences', in *International Conference on Social Robotics*, pp. 52–61, Amsterdam, Netherlands, (2011).

[12] R. Mead, A. Atrash, and M. J. Matarić, 'Recognition of spatial dynamics for predicting social interaction', in *6th ACM/IEEE International Conference on Human-Robot Interaction*, pp. 201–202, Lausanne, Switzerland, (2011).

[13] R. Mead, A. Atrash, and M. J. Matarić, 'Representations of proxemic behavior for human-machine interaction', in *NordiCHI 2012 Workshop on Proxemics in Human-Computer Interaction*, NordiCHI'12, Copenhagen, Denmark, (2012).

[14] R. Mead, A. Atrash, and M. J. Matarić, 'Automated proxemic feature extraction and behavior recognition: Applications in human-robot interaction', *International Journal of Social Robotics*, **5**(3), 367–378, (2013).

[15] R. Mead and M. J. Matarić, 'A probabilistic framework for autonomous proxemic control in situated and mobile human-robot interaction', in *7th ACM/IEEE International Conference on Human-Robot Interaction*, HRI'12, pp. 193–194, Boston, Massachusetts, (2012).

[16] R. Mead and M. J. Matarić, 'Perceptual models of human-robot proxemics', in *14th International Symposium on Experimental Robotics*, ISER'14, p. to appear, Marrakech/Essaouira, Morocco, (2014).

[17] R. Mead, E. Wade, P. Johnson, A. St. Clair, S. Chen, and M. J. Mataric., 'An architecture for rehabilitation task practice in socially assistive human-robot interaction', in *Robot and Human Interactive Communication*, pp. 404–409, (2010).

[18] J. Mumm and B. Mutlu, 'Human-robot proxemics: Physical and psychological distancing in human-robot interaction', in *6th ACM/IEEE International Conference on Human-Robot Interaction*, HRI-2011, pp. 331–338, Lausanne, (2011).

[19] A Murata, 'Basic characteristics of human's distance estimation', in *1999 IEEE International Conference on Systems, Man, and Cybernetics*, volume 2 of *SMC'99*, pp. 38–43, (1999).

[20] J. Nelder and R. Wedderburn, 'Generalized linear models', *Journal of the Royal Statistical Society*, **135**(3), 370–384, (1972).

[21] I. Rae, L. Takayama, and B. Mutlu, 'The influence of height in robot-mediated communication', in *8th ACM/IEEE International Conference on Human-Robot Interaction*, HRI-2013, pp. 1–8, Tokyo, Japan, (2005).

[22] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita, 'How to approach humans?: Strategies for social robots to initiate interaction', in *HRI*, pp. 109–116, (2009).

[23] S.S. Stevens, 'On the psychological law', *Psychological Review*, **64**, 153–181, (2007).

[24] L. Takayama and C. Pantofaru, 'Influences on proxemic behaviors in human-robot interaction', in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, IROS'09, pp. 5495–5502, (2009).

[25] A. Tapus, M.J. Matarić, and B. Scassellati, 'The grand challenges in socially assistive robotics', *IEEE Robotics and Automation Magazine*, **14**(1), 35–42, (2007).

[26] Elena Torta, Raymond H. Cuijpers, and James F. Juola, 'Design of a parametric model of personal space for robotic social navigation', *International Journal of Social Robotics*, **5**(3), 357–365, (2013).

[27] Elena Torta, Raymond H. Cuijpers, James F. Juola, and David van der Pol, 'Design of robust robotic proxemic behaviour', in *Proceedings of the Third International Conference on Social Robotics*, ICSR'11, pp. 21–30, (2011).

[28] M.L. Walters, K. Dautenhahn, R.T. Boekhorst, K.L. Koay, D.S. Syrdal, and C.L.: Nehaniv, 'An empirical framework for human-robot proxemics', in *New Frontiers in Human-Robot Interaction*, pp. 144–149, Edinburgh, (2009).

[29] M.L. Walters, D.S. Syrdal, K.L. Koay, K. Dautenhahn, and R. te Boekhorst, 'Human approach distances to a mechanical-looking robot with different robot voice styles', in *The 17th IEEE International Symposium on Robot and Human Interactive Communication*, RO-MAN 2008, pp. 707–712, (2008).