What's The Point? Tradeoffs Between Effectiveness and Social Perception When Using Mixed Reality to Enhance Gesturally Limited Robots

Jared Hamilton*
jghamilton@mines.edu
MIRRORLab
Colorado School of Mines
Golden, Colorado

Thao Phung*
thaophung@mines.edu
MIRRORLab
Colorado School of Mines
Golden, Colorado

Nhan Tran
nttran@mines.edu
MIRRORLab
Colorado School of Mines
Golden, Colorado

Tom Williams
twilliams@mines.edu
MIRRORLab
Colorado School of Mines
Golden, Colorado

ABSTRACT

Mixed Reality visualizations provide a powerful new approach for enabling gestural capabilities on non-humanoid robots. This paper explores two different categories of mixed-reality deictic gestures for armless robots: a virtual arrow positioned over a target referent (a non-ego-sensitive allocentric gesture) and a virtual arm positioned over the gesturing robot (an ego-sensitive allocentric gesture). Specifically, we present the results of a within-subjects Mixed Reality HRI experiment (N=23) exploring the trade-offs between these two types of gestures with respect to both objective performance and subjective social perceptions. Our results show a clear trade-off between performance and social perception, with non-ego-sensitive allocentric gestures enabling faster reaction time and higher accuracy, but ego-sensitive gestures enabling higher perceived social presence, anthropomorphism, and likability.

KEYWORDS

Augmented Reality, Mixed Reality, Deictic Gesture, Social Presence, Anthropomorphism, Human-Robot Interaction

ACM Reference Format:

Jared Hamilton, Thao Phung, Nhan Tran, and Tom Williams. 2021. What's The Point? Tradeoffs Between Effectiveness and Social Perception When Using Mixed Reality to Enhance Gesturally Limited Robots. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21), March 8–11, 2021, Boulder, CO, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3434073.3444676

*The first two authors contributed equally to this paper.

HRI '21, March 8–11, 2021, Boulder, CO, USA

ACM ISBN 978-1-4503-8289-2/21/03…$15.00
https://doi.org/10.1145/3434073.3444676

1 INTRODUCTION

For robots to communicate effectively with humans, they must be capable of natural, human-like human-robot dialogue [9, 42, 56]. And, in contrast to dialogue agents and chatbots, interactive robots must be able to communicate with sensitivity to situated context [42, 62]. This requires three broad competencies: *environmental context sensitivity* (sensitivity to the spatially situated, large-scale, uncertain, and incompletely known nature of task environments [73]); *cognitive context sensitivity* (sensitivity to the working memory and attentional constraints of teammates [74]); and *social context sensitivity* (sensitivity to the relational context into which they are embedded, and the importance of strengthening and maintaining social relationships through adherence to social and moral norms [39, 76] and building of trust and rapport [19, 30, 50]).

For these three competencies to be mastered, robots must be able to understand and generate both verbal behaviors and nonverbal behaviors such as gesture and eye gaze. Nonverbal behaviors are critical for situated interaction [2, 18, 28, 47], and are integrally related to these three competencies. Deictic gestures such as pointing leverage environmental context by identifying nearby referents, especially when such referents are not currently known or attended to by interlocutors. These gestures are often generated due to cognitive context, to direct interlocutor attention [40] and reduce memory costs that would be otherwise imposed by communication [16, 47]. And gestures are often generated with sensitivity to social context, by mimicing the gestures of interlocutors to increase engagement and build rapport through mirroring [8]. While there has been a host of research on nonverbal behavior generation in HRI [1, 6, 7, 49, 52–54], not all robots are able to leverage the techniques developed in that prior work, as many robotic platforms lack the arms, heads, and eyes needed to generate expressive cues. This is especially true for mobile bases such as those used in warehouses, and free-flying drone platforms. While these types of robots may not be designed to be sociable, they still need gaze- and gestural-capabilities for situated communication. Accordingly, researchers have been investigating new methods for nonverbal signalling (e.g., directed lighting cues) that may achieve those goals typically addressed by physical gaze and gesture [12, 60].

Mixed-reality technologies such as the Microsoft HoloLens stand to enable exciting new approaches to human-robot interaction [59], including generation of nonverbal cues in this vein for robots with non-humanoid morphologies [77]. The space of visualizations used as mixed-reality deictic gestures (which can altogether be classified as *view-augmenting* Mixed Reality Interaction Design Elements in the Reality-Virtuality Interaction Cube framework of Williams, Szafir, and Chakraborti [75]) can be divided into at least four primary classes: allocentric gestures (e.g., circling a target referent in a user's Mixed Reality head-mounted display (MR-HMD)), perspective-free gestures (e.g., projecting a circle around a target referent on the floor of the shared environment), ego-sensitive allocentric gestures (e.g., pointing to a target referent using a simulated arm rendered in a user's MR HMD), and ego-sensitive perspective-free gestures (e.g., projecting a line from the robot to its target on the floor of the shared environment [77]. In previous work, Williams et al. [68] specifically investigated the first of these categories, allocentric gestures, and demonstrated that mixed-reality gestures can significantly increase the communicative effectiveness of non-humanoid robots [63, 68, 70].

One downside of these previous explorations of allocentric gesture is low ecological validity of evaluation context, with crowdworkers viewing interactive videos simulating the expected appearance of such gestures. Accordingly, participants in previous experiments had full Field of View and viewed the entire experimental environment through an unchanging vantage point. In realistic task contexts, users are unlikely to be able to view their entire task environment from a single perspective, and Mixed Reality deictic gestures must be delivered through platforms like the HoloLens, limiting the portion of the environment in which gestures can be displayed. In even moderately larger task contexts, these factors could result in users completely directing their Field of View and attention towards the regions where mixed-reality deictic gestures are being displayed, completely avoiding the non-humanoid robot generating the visualizations. This lack of attention towards the robot could have detrimental long-term effects on human-robot teaming, such as decreased trust, rapport, and situation awareness.

These challenges may be addressable by another form of Mixed Reality deictic gesture highlighted in Williams et al. [77]'s taxonomy: ego-sensitive allocentric gestures, in which simulated arms are rendered above the robot, and used to point just as physical arms would [29]. The use of such arms could increase the robot's anthropomorphism, and because users would need to consistently look towards the robot to see where it is pointing, such arms could increase the robot's social presence.

On the other hand, ego-sensitive allocentric gestures may come with their own challenges. Because users will need to follow the vector along which the robot is pointing and estimate which objects fall within the robot's deictic cone, they may be less accurate and efficient at determining the targets of those gestures, especially when target referents are far from the robot (the very context in which ego-sensitive allocentric gestures are expected to provide social benefits). In this paper, we systematically evaluate these expected differences in social- and task-oriented benefits between ego-sensitive and non-ego-sensitive forms of allocentric gesture.

2 RELATED WORK

2.1 Human Deictic Gesture

Deixis is a key component of human-human communication [37, 44]. Humans begin pointing while speaking even from infancy, using deictic gestures around 9-12 months [4], and mastering deictic reference around age 4 [14]. Among adults, deictic gestures remain a critical component of situated communication, helping direct interlocutor attention in order to establish joint and shared attention [2]. Deictic gestures also help humans express their thoughts, especially in environments in which verbal communication would be difficult such as in noisy factory environments [31]. Accordingly, HRI researchers have been investigating how to enable effective robotic understanding and generation of deictic gestures.

2.2 Robot Deictic Gesture

Widespread evidence has been found in the HRI literature for the effectiveness of robots' use of nonverbal cues, including deictic gestures such as *pointing*, across a variety of different contexts, including tabletop environments [53] and free-form direction-giving [46]. Not only can robots use deictic gestures just as effectively as humans for the purpose of shifting interlocutor attention [7], but moreover it has been shown that robots' use of deictic gestures also improves subsequent human recall and human-robot rapport [6].

Related to our work, Sauppé and Mutlu [54] investigated a group of robotic deictic gestures: touching, presenting, grouping, pointing, sweeping, and exhibiting. This group of gestures was inspired by [15], who studied human deictic gestures and concluded that humans use more than just pointing as deictic gestures. Sauppé and Mutlu examined the objective and subjective differences between these six categories deictic gestures; Similarly, in this work we examine the objective and subjective differences between two categories of *Mixed Reality* deictic gestures.

2.3 Mixed Reality Human-Robot Communication

While the utility of robot deictic gestures has been well demonstrated, the generation of these gestures by robots is subject to a number of constraints. First, the ability to generate precise and natural deictic gestures typically requires robot arms: a morphological choice that is not only often prohibitively expensive, but moreover does not make sense for all use cases due to factors such as arm size and weight (e.g., in the case of drones). In order to enable gestural capabilities while avoiding the financial and morphological limitations of traditional deictic gesture, some researchers have recently been investigating the use of *Mixed Reality* gestures that could serve these same purposes.

Mixed Reality provides opportunities for a host of new forms of human-robot communication across domains such as task-driven interaction [26, 27, 57], collaboration exploration [48], and social interaction [13, 65]. For example, Frank et al. [26] presents an MR interface that shows a robot's reachable spatial regions in green and unreachable regions in red so that humans can better understand where and how to pass objects to the robot. In [27], humans and robots work together to assemble and move a car door. In the MR interface, the robot gives the human teammate information about

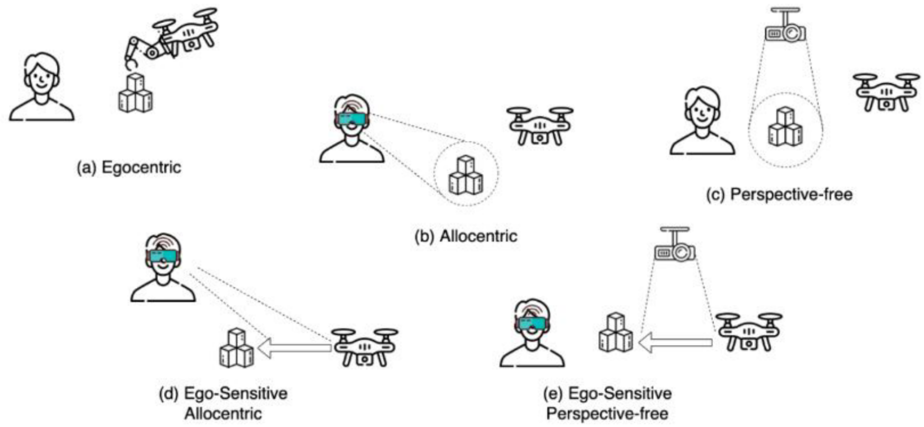


Figure 1: Categories of mixed-reality gestures [69] (used with permission).

its working area, what part of the door it will work on next, moving instructions, and the success of each subtask. In [13], the robot projects its trajectory and the spatial region that it would occupy while moving on the floor so the nearby humans can see and make a move to avoid collision with the robot.

Even for referential communication alone, Augmented and Mixed Reality afford many new forms of communication. For example, Sibirtseva et al. [57] presents a Mixed Reality interface in which circles and labels are used to indicate the set of objects a robot is considering during reference resolution. By looking at those candidate objects, humans can provide additional explanation to the robot to assist disambiguation. In contrast, Reardon et al. [48] present a MR interface for collaborative human-robot exploration tasks. Once a robot finds a target object, a trajectory from the human's current location to that object's location is visualized in the MR interface, to aid navigation to that object.

While Sibirtseva et al. focus on *passive backchannel communication* and Reardon et al. focus on *passive nonlinguistic communication*, we are particularly interested in *active linguistic communication* in which Mixed Reality imagery can be visualized alongside spoken language as an alternative to physical gesture. In previous work, Williams et al. [69] presented a framework for categorizing the different types of Mixed Reality deictic gestures that could be used in this way, delineating between four main categories of deictic gestural cues unique to Mixed Reality environments (Fig. 1):

(1) **Perspective-free gestures:** gestures that are projected onto the environment from a third-party perspective. Weng et al. [67], for example, studied where a robot should project arrows onto a tabletop to reference target objects.
(2) **Ego-sensitive perspective-free gestures:** gestures that are projected onto an environment in a way that connects the speaker (i.e., the robot) to its referent, e.g., if Weng et al. had generated their visualizations in such a way that the base of generated arrows originated at the robot.
(3) **Allocentric gestures:** gestures that pick out the speaker's target referent using imagery generated from the viewer's perspective (e.g., within MR-HMDs). Williams et al. [68, 70] prototyped one category of Mixed Reality deictic gesture, non-ego-sensitive allocentric gestures (e.g., gestures like circles and arrows, generated from a user's perspective, without taking the robot generator into account), and provided the first evidence for the effectiveness of those gestures within an online evaluation testbed. More recently, [63] demonstrated the effectiveness of these for the first time on realistic robotic and mixed-reality hardware.
(4) **Ego-sensitive allocentric gestures:** gestures that connect the speaker to its referent within the viewer's perspective. For example, [29] augment an otherwise armless robot with virtual arms shown in an MR interface.

However, while researchers are beginning to prototype and explore gestures within each of these categories, there has been no previous research comparing gestures between these categories. In this work, we present the first such research, systematically comparing between the non-ego-sensitive and ego-sensitive categories of allocentric Mixed Reality deictic gestures. We choose to investigate these two categories in particular because we believe they present a challenging tradeoff between objective task performance and subjective robotic perception, as detailed in the next section.

2.4 Social Perceptions of Robots

While most research on robot deictic gesture has evaluated gestures based on *objective*, task-driven metrics such as accuracy and reaction time of gesture interpretation, recent work has also sought to evaluate how those gestures are *subjectively* perceived. Sauppé and Mutlu [54], for example, evaluate gestures on the basis of how *natural* they appear to be, and Williams et al. [69, 71] evaluate gestures on the basis of their impact on robot *likability*. We believe that the two categories of gestures we examine in this work may present a challenging design case in which each of the two gestural options optimizes a fundamentally different category of metric.

First, we would expect non-ego-sensitive allocentric gestures (e.g., circles and arrows) to perform better on the objective measures delineated above. When a robot uses ego-sensitive allocentric gesures (e.g., virtual arms attached to its body), viewers must follow the deictic cone extending from the robot along its arm and attempt to determine which objects might fall within that cone. In contrast, when a robot uses non-ego-sensitive deictic gestures (e.g., circles and arrows), the robot's intended target is immediately and obviously picked out in the user's Field of View, providing little opportunity for inaccuracy or inefficiency.

However, we expect ego-sensitive allocentric gestures (e.g., virtual arms) may in turn perform better on *subjective* measures, such as anthropomorphism, social presence, likability, warmth, and competence. Below, we will examine each of these categories and articulate why we believe the use of ego-sensitive allocentric gestures may lead to higher subjective ratings in those categories.

Anthropomorphism is the projection of human characteristics to non-human entities [20, 21, 23]. Within HRI, researchers have framed anthropomorphism in terms of the contrast between robots designed in the image of humans (with humanlike features and appendages) and robots designed in the image of animals (i.e., zoomorphism) or robots with purely functional designs [25]. Anthropomorphism has been shown to be a valuable for robot design as it cues models of human-human interaction, facilitating sensemaking and mental model alignment [20], leading humans to be more willing to interact, accept, and understand robot's behaviors, and reducing human stress during interaction [36]. Moreover, robots that use gesture have been found to appear more anthropomorphic [52]. We expect robots using ego-sensitive allocentric gestures to be viewed as more anthropomorphic because they can provide human-like morphological features to otherwise mechanomorphic robots, can provide the illusion of motion and life to otherwise inert robots, and are more directly analogous to traditional physical robot gestures.

Social Presence is the feeling of being in the company of another social actor [58], and has been long explored within media studies due to the potential for a technology's social presence to enable more effective social and group interactions [5, 38]. Within HRI, researchers have found that robot social presence facilitates user enjoyment and desire to re-interact [33], perhaps due to our innate drive to seek out, engage in, and respond to socially interactive behaviors with other social actors. Social presence is also related to anthropomorphism in interesting ways. Specifically, Nowak and Biocca [45] found that very low and very high levels of anthropomorphism led to lower levels of social presence than middling levels of anthropomorphism. We predict that robots using ego-sensitive allocentric gestures will be viewed as more anthropomorphic, but not *highly* anthropomorphic, both due to their status as obvious augmentations and because the arms we explore in this work are not highly photorealistic renderings of human arms. We further predict that this will lead robots to be perceived as having greater social presence. Moreover, we believe that the use of robot arms is likely to engender greater social presence due to the impact these arms may have on visual attention. That is, while cues like circles and arrows may be interpreted without looking at or considering the robot generating them, virtual arms require the user to repeatedly regard the robot in order to interpret its gestures. We believe this is likely to significantly increase the perceived presence of the robot.

Warmth and Competence are social psychological constructs that are at the core of social judgment and are nearly entirely responsible for social perceptions among humans [24]. While warmth captures whether an actor is sociable and well-intentioned, the competence captures whether they have potential to deliver on those intentions. Warmth and competence are thus valuable within human-human interaction as they lead to more positive emotions [24]. Within HRI, warmth and competence have been found to be key predictors for human preferences between robots and robot behaviors [55], and have been shown to lead to more positive human-robot interactions [10]. These concepts are also related to those discussed above: warmth in particular is often associated with Social Presence [32], and anthropomorphism has been shown in certain contexts to directly lead to greater warmth [35] and competence-based trust [66]. Because we predict that robots using ego-sensitive allocentric gestures will be viewed as more anthropomorphic and as having higher social presence, we thus believe that this will then also lead them to be perceived as more warm and competent.

Likability is a key usability metric used to summarize peoples' overall perceptions of technology, and has been one of the primary metrics used across the HRI field [3]. In gesture-related contexts, including in Mixed Reality contexts, it has been found that gesture use can lead to increased likability [52, 70, 76]. Because we predict that robots using ego-sensitive allocentric gestures will be viewed as more anthropomorphic, and as having higher social presence, warmth, and competence, we thus believe that this will then also lead them to be perceived as more likeable.

2.5 Hypotheses

Based on our review of the previous work discussed above, we formulate the following concrete hypotheses:

H1:A robot that uses non-ego-sensitive allocentric gestures (i.e., arrows drawn over target referents) when referring to target referents will (**H1.1**) be *more effective* than a robot using ego-sensitive allocentric gestures (i.e., pointing using virtual arms) as measured by (1) accuracy and (2) reaction time, and (**H1.2:**) these benefits would be more pronounced for objects farther away from the robot.

H2:A robot that uses non-ego-sensitive allocentric gestures (i.e., arrows drawn over target referents) when referring to target referents will: (**H2.1:**) have *lower social perception* than a robot using ego-sensitive allocentric gestures (i.e., pointing using virtual arms) as measured by (1) social presence, (2) anthropomorphism, (3) likability, (4) warmth, and (5) perceived competence, and (**H2.2:**) these detriments would be more pronounced for objects farther away from the robot.

3 EXPERIMENT

To investigate these hypotheses, we conducted a within-subjects human-subject experiment in which participants interacted with a robot in a mixed-reality HRI context.

3.1 Experimental Design

Our experiment consisted of a series of four Latin-Square order-counterbalanced experiment blocks in each of which participants interacted with a robot in a mixed-reality HRI context. In each of these experiment blocks, participants performed a *gesture understanding* task consisting of ten trials. No time limit was imposed, but participants were encouraged to complete the task as quickly as possible. In each trial, participants' robot teammate gestured to one

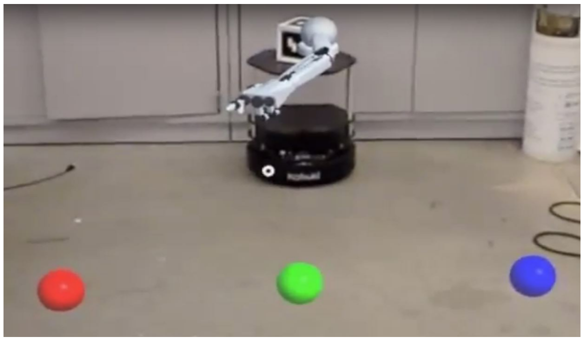


Figure 2: Robot arm gesturing to holographic sphere (Not in experimental environment).

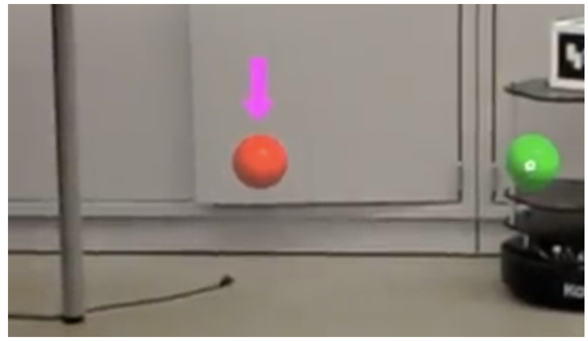


Figure 3: Virtual arrow pointing to holographic sphere (the robot arm doesn't show up in this case)

of three target referents (multi-colored spheres) at random using a Mixed Reality deictic gesture, and participants were required to "click" (using an air-tap gesture) on the referent they believed the robot was gesturing towards. Between experiment blocks, participants completed surveys assessing their perceptions of the robot and its gestures, as described in Sec. 3.4.

Experimental blocks differed according to a 2x2 design in which two independent variables were manipulated: *Gesture Type* and *Referent Distance*. Specifically, each task was conducted in one of two *Gesture Type* conditions: in two of the four within-subject blocks, participants interacted with a robot that gestured using *ego-sensitive allocentric gestures* in which a virtual arm reached out and pointed towards target referents; in the other two within-subject blocks, participants interacted with a robot that gestured using *non-ego-sensitive allocentric gestures* in which an arrow appeared over target referents. Each of these two conditions was then further subdivided into two *Referent Distance* conditions: a *robot-close* condition in which the robot's target referents were approximately one meter from the robot and two meters from the human, and a *robot-distant* condition, in which the robot's target referents were approximately two meters from the robot and one meter from the human.

3.2 Experimental Apparatus

Robotic Platform: As shown in Fig. 2, a Kabuki TurtleBot 2 was used, affixed with an MR cube: a 12cm cardboard cube with fiducial markers on each face. This cube served as an anchor for the robot arm in the *arm* conditions, and allowed the HoloLens to determine the robot's position in all conditions.

Mixed-Reality Head-Mounted Display: The MR-HMD used in this experiment was a Microsoft HoloLens, a commercial-grade stereographic Mixed-Reality Headset with a $30° \times 17.5°$ Field of View. Participants' air-tap gestures were detected using the HoloLens' built-in gesture recognition capabilities.

Mixed-Reality Deictic Gestures: Two mixed-reality deictic gestures were designed in Unity; a *arrow* and *arm*. The arrow was a simple magenta arrow that statically appeared over target referents, shown in Fig. 3. The arm used a virtual arm model created and textured using Blender, and animated in Unity using a custom-built key-frame-based animation library.

Experimental Application: The experimental procedure and autonomous robot behavior were coded as a Unity application deployed onto the HoloLens. When viewing the scene through the HoloLens, participants perceived three spheres (red, green, and blue) hovering a half-meter above the ground between the subject and the TurtleBot. Different colors were used for compatibility with intended future work. The HoloLens also enabled participants to see the robot's gestures. In the *arm* conditions, an arm was always visible over the TurtleBot. In the *arrow* conditions, the arm was invisible and an arrow instead appeared over target referents.

3.3 Procedure

Upon arriving at the lab, participants provided informed consent and completed a demographic survey. Participants were then introduced to the TurtleBot, the HoloLens, and the task through both verbal instruction and an interactive tutorial designed in Unity and deployed on the HoloLens. The tutorial interface showed instruction text and virtual red, blue, green spheres, walking participants through a sample experimental trial. During the tutorial, the participants learned how to use air-tap gestures to choose a sphere. After demonstrating the ability to successfully air-tap a target sphere three times, participants proceeded to the experiment. After completing the experiment, participants were paid and debriefed.

3.4 Measures

To assess our two hypotheses, seven key metrics were collected during our experiment, including two objective measures and five subjective measures.

Objective Measures

Our first hypothesis was assessed using two objective measures:
Accuracy was measured as the percent of trials in which the target selected after a gesture was in fact the target of that gesture.
Reaction Time was measured as the time (in seconds) from the time a gesture was triggered to the time a user selected the object they believed to be indicated by that gesture.

Subjective Measures

Our second hypothesis was assessed using five sets of survey questions administered after each experiment block. Each set of survey questions was a Likert scale comprised of 5-6 items asking for statement agreement or disagreement on a 1-5 scale.

Social Presence was measured using the Almere Social Presence scale [34].
Anthropomorphism was measured using the Godspeed II Anthropomorphism scale [3].
likability was measured using the Godspeed II likability scale [3].
Warmth was measured using the RoSAS Warmth scale [11].
Competence was measured using the RoSAS Competence scale [11].

3.5 Participants

24 participants were recruited from the Colorado School of Mines through web postings and flyers (14 male, 10 female) for an ethics board approved experiment. Participants ranged in age from 18 to 52 (M=22.46, SD=7.86). 20 of the 24 participants had not previously engaged in any experiments from our laboratory involving Mixed Reality. One participant failed to complete the experiment, leaving 23 usable data points.

3.6 Analysis

Data analysis was performed within a Bayesian analysis framework using the JASP 0.8.5.1 [61] software package, using the default settings justified by Wagenmakers et al. [64]. For each measure, a Bayesian Repeated Measures Analysis of Variance (RM-ANOVA) [17, 43, 51] was performed, using Gesture Type and Referent Distance as random factors. Bayes Inclusion Factors Across Matched Models ("Baws Factors") [41] were then computed for each candidate main effect and interaction, indicating (in the form of a Bayes Factor) the evidence weight of all candidate models including that effect compared to the evidence weight of all candidate models not including that effect. Analysis of Likert Scale data was performed after averaging responses within each scale.

4 RESULTS

4.1 Objective Results

Figure 4 summarizes our main objective results.

Accuracy – Our results provided strong evidence in favor of an effect of Gesture Type on accuracy (Bf 16.376), as shown in Fig. 4a, suggesting that our data were 16 times more likely to be generated under models in which Gesture Type was included than under those in which it was not, and specifically that when virtual arrows were used, participants were more accurate in determining the intended target of those gestures (close distance (M=1, SD=0), far distance (M=0.996, SD=0.021)) than when virtual arms were used (close distance (M=0.935, SD=0.204), far distance (M=0.891, SD=0.195)). However, anecdotal evidence was found *against* an interaction effect between Gesture Type and Referent Distance on accuracy (BF 0.415), with the data 1/0.415 = 2.41 times less likely to have been generated under models including such an interaction.

Reaction time – Our results provided strong evidence in favor of an effect of Gesture Type on reaction time (Bf 22.264), as shown in Fig. 4b, suggesting specifically that when virtual arrows were used, participants could more quickly identify the targets of those gestures (close distance (M=2.265, SD=1.047), far distance (M=2.139, SD=0.757) than when virtual arm were used (close distance (M=3.678, SD=2.941), far distance (M=5.152, SD=6.154)). However, anecdotal evidence was found *against* an interaction effect between Gesture Type and Referent Distance on reaction time (BF 0.455), that is, the data was 1/0.455 = 2.198 times less likely to have been generated under models including such an interaction.

4.2 Subjective Results

Figure 5 summarizes our main subjective results.

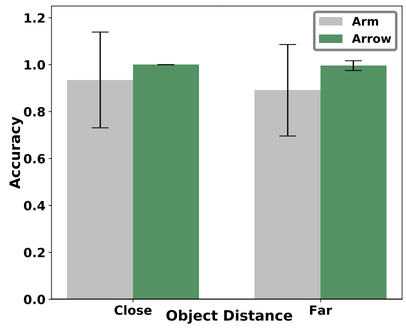
Social Presence – Our results provided extreme evidence in favor of an effect of Gesture Type on social presence (Bf 440.332), as shown in Fig. 5c, suggesting specifically that when virtual arrows were used, participants viewed the robot as having lower social presence (close distance (M=9.458, SD=3.845), far distance (M=10.125, SD=3.327)) than when virtual arms were used (close distance (M=11.792, SD=2.570), far distance (M=11.250, SD=3.179)). However, our results provided no significant evidence for or against of an interaction between Gesture Type and Referent Distance on social presence, suggesting that more data must be collected before a conclusion can be reached. As shown in Fig. 5c, it is plausible but not yet verifiable that when objects were *close* to the robot that use of virtual arms led to greater robotic social presence.

Anthropomorphism – Our results provided strong evidence in favor of an effect of Gesture Type on anthropomorphism (Bf 6026.6), as shown in Fig. 5a, suggesting specifically that when virtual arrows were used, participants viewed the robot as having lower anthropomorphism (close distance (M=9.833, SD=4.239), far distance (M=9.667, SD=4.239)) than when virtual arms were used (close distance (M=12.125, SD=3.826), far distance (M=12.417, SD=4.671)). However, moderate evidence was found *against* an interaction effect between Gesture Type and Referent Distance on perceived anthropomorphism (BF 0.301), that is, the data was 1/0.301 = 3.322 times less likely to have been generated under models including such an interaction.

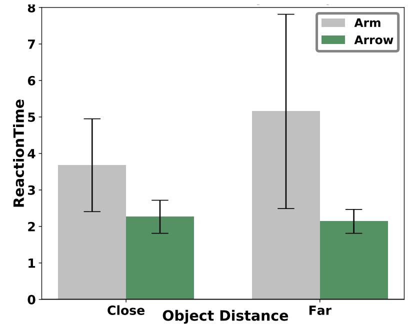
Likability – Our results provided moderate evidence in favor of an effect of Gesture Type on likability (Bf 6.145), as shown in Fig. 5b, suggesting specifically that when virtual arrows were used, participants viewed the robot as having lower likability (close distance (M=15.500,SD=3.845), far distance (M=15.583, SD=3.263)) than when virtual arms were used (close distance (M=16.917,SD=3.855), far distance (M=16.583, SD=3.764)). However, moderate evidence was found *against* an interaction effect between Gesture Type and Referent Distance on perceived likability (BF 0.319), that is, the data was 1/0.319 = 3.13 times less likely to have been generated under models including such an interaction.

Warmth – Our results provided no significant evidence for or against of an effect of Gesture Type on warmth (Bf 1.567), as shown in Fig. 5d, suggesting that more data must be collected before a conclusion can be reached. Moreover, moderate evidence was found *against* an interaction effect between Gesture Type and Referent Distance on perceived warmth (BF 0.328), that is, the data was 1/0.328 = 3.049 times less likely to have been generated under models including such an interaction.

Competence – Our results provided no significant evidence for or against of an effect of Gesture Type on competence (Bf 1.194), as shown in Fig. 5e, suggesting that more data must be collected before a conclusion can be reached. Moreover, moderate evidence was found *against* an interaction effect between Gesture Type and Referent Distance on perceived competence (BF 0.284), that is, the

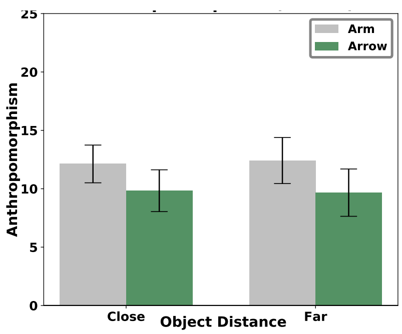


(a) Accuracy (95% CI)

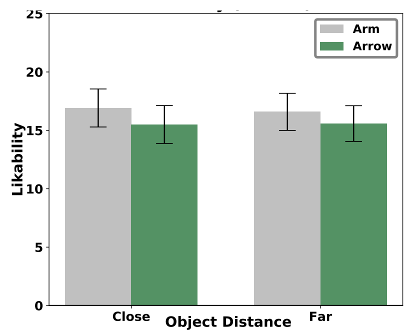


(b) Reaction Time (95% CI)

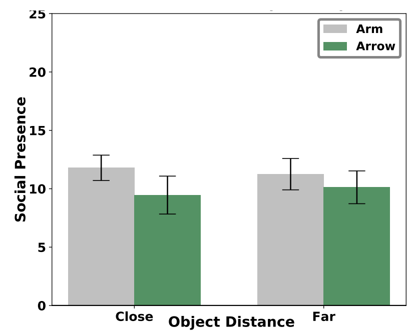
Figure 4: Objective Results



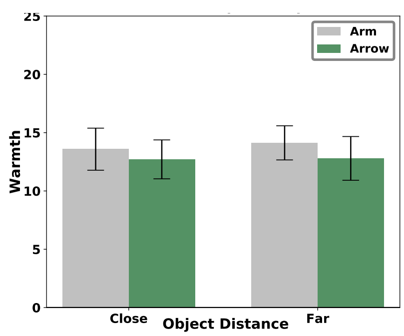
(a) Anthropomorphism (95% CI)



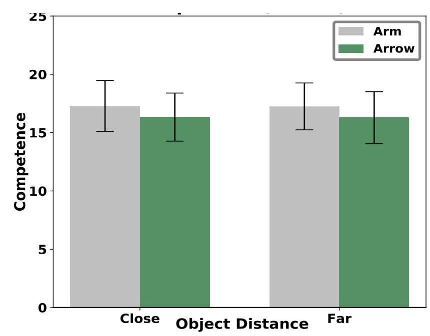
(b) Likability (95% CI)



(c) Social Presence (95% CI)



(d) Warmth (95% CI)



(e) Competence (95% CI)

Figure 5: Subjective Results

data was 1/0.284 = 3.521 times less likely to have been generated under models including such an interaction.

5 DISCUSSION

5.1 Hypothesis One

We hypothesized that a robot that uses virtual arrows when referring to target referents would: **(H1.1)** be more effective than a robot using virtual arms, as measured by accuracy and reaction time, and **(H1.2)** that these benefits would be more pronounced for objects farther away from the robot.

Our results support Hypothesis H1.1 but not Hypothesis H1.2. Our result suggests that a robot using virtual arrows is more effective than a robot using virtual arms: virtual arrows allowed users to complete the task faster and more accurately than virtual arms. This is unsurprising as virtual arrows directly pick out target referents without users needing to follow and interpret a deictic cone. While in the arrow scenario referent distance did not appear to impact accuracy and reaction time, in the arm scenario such an effect was observed: when the target referent was close to the robot, users could more accurately and quickly identify it.

While our Hypothesis H1.1 are supported, the results are inconclusive for Hypothesis H1.2. The anecdotal evidence against an interaction effect between Gesture Type and Referent Distance on accuracy (BF 0.415) and reaction time (BF 0.455) is not strong enough to conclusively rule out an effect, and visual inspection suggests there may indeed have been effects of distance on both accuracy and reaction time, in which task performance improved for virtual arms when referents were closer to the robot. More data will be needed to confirm or rule out these effects, in a larger environment allowing greater distinction between distance conditions.

5.2 Hypothesis Two

We hypothesized that a robot that uses virtual arrows when referring to target referents would: **(H2.1)** be have lower social perception than a robot using arms as measured by social presence, anthropomorphism, likability, warmth, and perceived competence, and **(H2.2)** that these detriments would be more pronounced for objects farther away from the robot. We will thus separately assess this hypothesis for each of these subjective measures.

Our results support Hypothesis H2.1 but fail to support Hypothesis H2.2. First, our results suggest that robots using arm have higher social perception in terms of anthropomophism, social presence, and likability than non-ego-sensitive allocentric gestures, which we believe is due to the human-like, animated morphology provided by virtual arms. Second, our results suggest that robots using arm were also perceived as more likable than robots using virtual arrows, which we believe is due to that anthropomorphism and social presence. Again, we believe that while virtual arms continually draw the user's visual attention back to the robot, when virtual arrows are used, users can essentially ignore the robot generating them without any loss in performance. These findings were also observed to be highly sensitive to distance. First, the robot using virtual arms perceived to have higher anthropomorphism when referring to objects closer to it, which we believe to be due to increased time with the animated robot in frame within the HoloLens' limited Field of View. Second, the robot using virtual arms was rated as more likability and more socially present when referring to objects farther from it; effects which are not yet clear how to interpret.

Finally, our results neither supported or refuted an effect of Gesture Type on warmth or competence. We expect that these findings may in part due to the *actual* increase in competence for robots that used virtual arrows. That is, the decreased anthropomorphism and social presence may have led these robots using arm to be perceived as more competent, but overall robots using those gestures were in fact overall *less* competent in picking out target referents than robots using virtual arrows.

5.3 Limitations and Future Work

The main limitation of our experiment is small sample size, which while necessary due to pandemic-related campus shutdowns, yielded unnecessarily inconclusive results in some analyses. Specifically, several analyses produced Bayes Factors between 1/3 and 3, suggesting inconclusive results neither supporting nor refuting our hypotheses, and instead suggesting the need to collect more data. While in the wake of COVID-19 many experiments are moving online [22], and while some preliminary MR-for-HRI experiments have indeed been conducted online [68], the nature of this experiment (especially with respect to physical head and eye motions to shift the Field of View of Mixed Reality) is not only ill-suited to online experimentation but would also benefit from measurement options available only in person.

In future work we hope to leverage the HoloLens 2, which has a larger Field of View, which would allow us to identify whether our observed effects were due in part to the need for participants to physically shift their overall gaze in order to keep visualizations in-frame. Moreover, the eye-tracking capabilities of the HoloLens 2 would allow us to determine whether there were in fact differences in gaze-towards-robot between different Mixed Reality gestures. Moreover, while the use of MR experiments increases Pandemic-related safety in some ways (e.g., by decreasing contact between humans and task-relevant objects [72]), it also necessitates the use of inherently high-contact equipment (i.e., the MR-HMD itself).

Finally, our results revealed an interesting design challenge, in which designers should use non-ego-sensitive allocentric gestures like circles and arrows if they wish to maximize short-term task performance, but should use ego-sensitive allocentric gestures like virtual arms if they wish to maximize social dimensions likely to impact long-term task performance. In future work, we hope to explore whether these gesture categories can be used in conjunction to achieve the best of both worlds, or whether this would cognitively and visually overload users.

6 CONCLUSION

We conducted an N=24 HRI experiment to compare two categories of mixed-reality deictic gestures for armless robots: a virtual arrow positioned over a target referent (a non-ego-sensitive allocentric gesture) and a virtual arm positioned over the robot (an ego-sensitive allocentric gesture). Our results suggest that non-ego-sensitive allocentric gestures enable faster reaction time and higher accuracy, while ego-sensitive gestures enable higher perceived social presence, anthropomorphism, and likability. This presents a clear design trade-off: our results suggest the need for different mixed reality gestures to be used in different application domains depending on the nature of the task and the intended relationship the designer seeks to establish between human and robot. Most domains in which Mixed Reality HRI is currently being envisioned are task-oriented domains, such as the use of collaborative robots in advanced manufacturing environments. In such domains, our results suggest that designers may wish to leverage non-ego-sensitive allocentric mixed reality gestures. On the other hand, for robots designed for more socially oriented domains, or even task-oriented domains where it is advantageous to highlight a robot's social or anthoropomorphic characteristics, ego-sensitive allocentric gestures may instead be preferable. Finally, our results highlight important connections between visual attention, anthropomorphism, social presence, warmth, and competence, which are critical to the design of interactive robots even beyond mixed reality environments.

ACKNOWLEDGMENTS

This work was funded in part by National Science Foundation grants IIS-1909864 and CNS-1810631.

REFERENCES

[1] Henny Admoni and Brian Scassellati. 2017. Social eye gaze in human-robot interaction: a review. *Journal of Human-Robot Interaction* 6, 1 (2017), 25–63.

[2] Adrian Bangerter and Max M Louwerse. 2005. Focusing attention with deictic gestures and linguistic expressions. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 27.

[3] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics* 1, 1 (2009), 71–81.

[4] Elizabeth Bates. 1976. *Language and context: The acquisition of pragmatics.* Academic Press.

[5] Frank Biocca, Chad Harms, and Judee K Burgoon. 2003. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & virtual environments* 12, 5 (2003), 456–480.

[6] Cynthia Breazeal, Cory D Kidd, Andrea Lockerd Thomaz, Guy Hoffman, and Matt Berlin. 2005. Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *2005 IEEE/RSJ international conference on intelligent robots and systems.* IEEE, 708–713.

[7] Andrew G Brooks and Cynthia Breazeal. 2006. Working with robots and objects: Revisiting deictic reference for achieving spatial common ground. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction.* 297–304.

[8] Judee K Burgoon and Gregory D Hoobler. 1994. Nonverbal signals. *Handbook of interpersonal communication* 2 (1994), 229–285.

[9] Rehj Cantrell, Paul Schermerhorn, and Matthias Scheutz. 2011. Learning actions from human-robot dialogues. In *2011 RO-MAN.* IEEE, 125–130.

[10] Colleen M Carpinella, Alisa B Wyman, Michael A Perez, and Steven J Stroessner. 2017. The robotic social attributes scale (RoSAS) development and validation. In *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction.* 254–262.

[11] Colleen M Carpinella, Alisa B Wyman, Michael A Perez, and Steven J Stroessner. 2017. The Robotic Social Attributes Scale (RoSAS) Development and Validation. In *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction.* 254–262.

[12] Elizabeth Cha, Yunkyung Kim, Terrence Fong, Maja J Mataric, et al. 2018. A survey of nonverbal signaling methods for non-humanoid robots. *Foundations and Trends® in Robotics* 6, 4 (2018), 211–323.

[13] Ravi Teja Chadalavada, Henrik Andreasson, Robert Krug, and Achim J Lilienthal. 2015. That's on my mind! robot to human intention communication through on-board projection on shared floor space. In *2015 European Conference on Mobile Robots (ECMR).* IEEE, 1–6.

[14] Eve V Clark and CJ Sengul. 1978. Strategies in the acquisition of deixis. *Journal of child language* 5, 3 (1978), 457–475.

[15] Herbert H Clark. 2005. Coordinating with each other in a material world. *Discourse studies* 7, 4-5 (2005), 507–525.

[16] Susan Wagner Cook, Terina Kuangyi Yip, and Susan Goldin-Meadow. 2012. Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and cognitive processes* 27, 4 (2012), 594–610.

[17] Martin J Crowder. 2017. *Analysis of repeated measures.* Routledge.

[18] Antonella De Angeli, Walter Gerbino, Giulia Cassano, and Daniela Petrelli. 1998. Visual display, pointing, and natural language: the power of multimodal interaction. In *Proceedings of the working conference on Advanced visual interfaces.* 164–173.

[19] Ewart J De Visser, Samuel S Monfort, Ryan McKendrick, Melissa AB Smith, Patrick E McKnight, Frank Krueger, and Raja Parasuraman. 2016. Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied* 22, 3 (2016), 331.

[20] Carl DiSalvo and Francine Gemperle. 2003. From seduction to fulfillment: the use of anthropomorphic form in design. In *Proceedings of the 2003 international conference on Designing pleasurable products and interfaces.* 67–72.

[21] Brian R Duffy. 2002. Anthropomorphism and robotics. *The society for the study of artificial intelligence and the simulation of behaviour* 20 (2002).

[22] David Feil-Seifer, Kerstin S Haring, Silvia Rossi, Alan R Wagner, and Tom Williams. 2020. Where to Next? The Impact of COVID-19 on Human-Robot Interaction Research. *ACM Transactions on Human-Robot Interaction (THRI)* 10, 1 (2020), 1–7.

[23] Julia Fink. 2012. Anthropomorphism and human likeness in the design of robots and human-robot interaction. In *International Conference on Social Robotics.* Springer, 199–208.

[24] Susan T Fiske, Amy JC Cuddy, and Peter Glick. 2007. Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences* 11, 2 (2007), 77–83.

[25] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A survey of socially interactive robots. *Robotics and autonomous systems* 42, 3-4 (2003), 143–166.

[26] Jared A Frank, Matthew Moorhead, and Vikram Kapila. 2017. Mobile mixed-reality interfaces that enhance human–robot interaction in shared spaces. *Frontiers in Robotics and AI* 4 (2017), 20.

[27] Ramsundar Kalpagam Ganesan, Yash K Rathore, Heather M Ross, and Heni Ben Amor. 2018. Better teaming through visual cues: how projecting imagery in a workspace can improve human-robot collaboration. *IEEE Robotics & Automation Magazine* 25, 2 (2018), 59–71.

[28] Susan Goldin-Meadow. 1999. The role of gesture in communication and thinking. *Trends in cognitive sciences* 3, 11 (1999), 419–429.

[29] Thomas Groechel, Zhonghao Shi, Roxanna Pakkar, and Maja J Matarić. 2019. Using Socially Expressive Mixed Reality Arms for Enhancing Low-Expressivity Robots. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN).* IEEE, 1–8.

[30] Peter A Hancock, Deborah R Billings, Kristin E Schaefer, Jessie YC Chen, Ewart J De Visser, and Raja Parasuraman. 2011. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors* 53, 5 (2011), 517–527.

[31] Simon Harrison. 2011. The creation and implementation of a gesture code for factory communication. In *GESPIN 2011: Gesture and speech in interaction.*

[32] Khaled Hassanein and Milena Head. 2007. Manipulating perceived social presence through the web interface and its impact on attitude towards online shopping. *International Journal of Human-Computer Studies* 65, 8 (2007), 689–708.

[33] Marcel Heerink, Ben Kröse, Vanessa Evers, and Bob Wielinga. 2010. Assessing acceptance of assistive social agent technology by older adults: the almere model. *International journal of social robotics* 2, 4 (2010), 361–375.

[34] Marcel Heerink, Ben Kröse, Vanessa Evers, and Bob Wielinga. 2010. Assessing acceptance of assistive social agent technology by older adults: the almere model. *International journal of social robotics* 2, 4 (2010), 361–375.

[35] Seo Young Kim, Bernd H Schmitt, and Nadia M Thalmann. 2019. Eliza in the uncanny valley: anthropomorphizing consumer robots increases their perceived warmth but decreases liking. *Marketing letters* 30, 1 (2019), 1–12.

[36] Dieta Kuchenbrandt, Nina Riether, and Friederike Eyssel. 2014. Does anthropomorphism reduce stress in HRI?. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction.* 218–219.

[37] Stephen C Levinson. 2004. 5 Deixis. *The handbook of pragmatics* (2004), 97.

[38] Matthew Lombard and Theresa Ditton. 1997. At the heart of it all: The concept of presence. *Journal of computer-mediated communication* 3, 2 (1997), JCMC321.

[39] Bertram F Malle and Matthias Scheutz. 2014. Moral competence in social robots. In *2014 IEEE international symposium on ethics in science, technology and engineering.* IEEE, 1–6.

[40] William Marslen-Wilson, Elena Levy, and Lorraine K Tyler. 1982. Producing interpretable discourse: The establishment and maintenance of reference. *Speech, place, and action* (1982), 339–378.

[41] S. Mathôt. 2017. Bayes like a Baws: Interpreting Bayesian repeated measures in JASP [Blog Post]. https://www.cogsci.nl/blog/interpreting-bayesian-repeated-measures-in-jasp.

[42] Nikolaos Mavridis. 2015. A review of verbal and non-verbal human–robot interactive communication. *Robotics and Autonomous Systems* 63 (2015), 22–35.

[43] RD Morey and JN Rouder. 2014. BayesFactor (Version 0.9. 9).

[44] Sigrid Norris. 2011. Three hierarchical positions of deictic gesture in relation to spoken language: a multimodal interaction analysis. *Visual Communication* 10, 2 (2011), 129–147.

[45] Kristine L Nowak and Frank Biocca. 2003. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 12, 5 (2003), 481–494.

[46] Yusuke Okuno, Takayuki Kanda, Michita Imai, Hiroshi Ishiguro, and Norihiro Hagita. 2009. Providing route directions: design of robot's utterance, gesture, and timing. In *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI).* IEEE, 53–60.

[47] Raedy Ping and Susan Goldin-Meadow. 2010. Gesturing saves cognitive resources when talking about nonpresent objects. *Cognitive Science* 34, 4 (2010), 602–619.

[48] Christopher Reardon, Kevin Lee, and Jonathan Fink. 2018. Come see this! augmented reality to enable human-robot cooperative search. In *2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR).* IEEE, 1–7.

[49] Laurel D Riek, Philip C Paul, and Peter Robinson. 2010. When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry. *Journal on Multimodal User Interfaces* 3, 1-2 (2010), 99–108.

[50] Paul Robinette, Ayanna M Howard, and Alan R Wagner. 2015. Timing is key for robot trust repair. In *International Conference on Social Robotics.* Springer, 574–583.

[51] Jeffrey N Rouder, Richard D Morey, Paul L Speckman, and Jordan M Province. 2012. Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology* 56, 5 (2012), 356–374.

[52] Maha Salem, Friederike Eyssel, Katharina Rohlfing, Stefan Kopp, and Frank Joublin. 2013. To err is human (-like): Effects of robot gesture on perceived anthropomorphism and likability. *International Journal of Social Robotics* 5, 3 (2013), 313–323.

[53] Maha Salem, Stefan Kopp, Ipke Wachsmuth, Katharina Rohlfing, and Frank Joublin. 2012. Generation and evaluation of communicative robot gesture. *International Journal of Social Robotics* 4, 2 (2012), 201–217.

[54] Allison Sauppé and Bilge Mutlu. 2014. Robot deictics: How gesture and context shape referential communication. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 342–349.

[55] Marcus M Scheunemann, Raymond H Cuijpers, and Christoph Salge. 2020. Warmth and Competence to Predict Human Preference of Robot Behavior in Physical Human-Robot Interaction. *arXiv preprint arXiv:2008.05799* (2020).

[56] Matthias Scheutz, Paul Schermerhorn, James Kramer, and David Anderson. 2007. First steps toward natural human-like HRI. *Autonomous Robots* 22, 4 (2007), 411–423.

[57] Elena Sibirtseva, Dimosthenis Kontogiorgos, Olov Nykvist, Hakan Karaoguz, Iolanda Leite, Joakim Gustafson, and Danica Kragic. 2018. A comparison of visualisation methods for disambiguating verbal requests in human-robot interaction. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 43–50.

[58] Paul Skalski and Ron Tamborini. 2007. The role of social presence in interactive agent-based persuasion. *Media psychology* 10, 3 (2007), 385–413.

[59] Daniel Szafir. 2019. Mediating human-robot interactions with virtual, augmented, and mixed reality. In *International Conference on Human-Computer Interaction*. Springer, 124–149.

[60] Daniel Szafir, Bilge Mutlu, and Terrence Fong. 2015. Communicating directionality in flying robots. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 19–26.

[61] JASP Team. 2018. JASP (Version 0.8.5.1)[Computer software].

[62] Stefanie Tellex, Nakul Gopalan, Hadas Kress-Gazit, and Cynthia Matuszek. 2020. Robots That Use Language. *Annual Review of Control, Robotics, and Autonomous Systems* 3 (2020).

[63] Nhan Tran, Kai Mizuno, Trevor Grant, Thao Phung, Leanne Hirshfield, and Tom Williams. 2019. Exploring Mixed Reality Robot Communication Under Different types of Mental Workload. In *Proceedings of the 3rd International Workshop on Virtual, Augmented, and Mixed Reality for HRI*.

[64] EJ Wagenmakers, J Love, M Marsman, T Jamil, A Ly, and J Verhagen. 2018. Bayesian inference for psychology, Part II: Example applications with JASP. *Psychonomic Bulletin and Review* 25, 1 (2018), 35–57.

[65] Atsushi Watanabe, Tetsushi Ikeda, Yoichi Morales, Kazuhiko Shinozawa, Takahiro Miyashita, and Norihiro Hagita. 2015. Communicating robotic navigational intentions. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 5763–5769.

[66] Adam Waytz, Joy Heafner, and Nicholas Epley. 2014. The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology* 52 (2014), 113–117.

[67] Thomas Weng, Leah Perlmutter, Stefanos Nikolaidis, Siddhartha Srinivasa, and Maya Cakmak. 2019. Robot Object Referencing through Legible Situated Projections. In *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 8004–8010.

[68] Tom Williams, Matthew Bussing, Sebastian Cabrol, Elizabeth Boyle, and Nhan Tran. 2019. Mixed Reality Deictic Gesture for Multi-Modal Robot Communication. In *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction*.

[69] Tom Williams, Matthew Bussing, Sebastian Cabrol, Elizabeth Boyle, and Nhan Tran. 2019. Mixed reality deictic gesture for multi-modal robot communication. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 191–201.

[70] Tom Williams, Matthew Bussing, Sebastian Cabrol, Ian Lau, Elizabeth Boyle, and Nhan Tran. 2019. Investigating the Potential Effectiveness of Allocentric Mixed Reality Deictic Gesture. In *Proceedings of the 11th International Conference on Virtual, Augmented, and Mixed Reality*.

[71] Tom Williams, Matthew Bussing, Sebastian Cabrol, Ian Lau, Elizabeth Boyle, and Nhan Tran. 2019. Investigating the potential effectiveness of allocentric mixed reality deictic gesture. In *International Conference on Human-Computer Interaction*. Springer, 178–198.

[72] Tom Williams, Leanne Hirshfield, Nhan Tran, Trevor Grant, and Nicholas Woodward. 2020. Using Augmented Reality to Better Study Human-Robot Interaction. In *International Conference on Human-Computer Interaction*. Springer, 643–654.

[73] Tom Williams and Matthias Scheutz. 2016. A Framework for Resolving Open-World Referential Expressions in Distributed Heterogeneous Knowledge Bases. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*.

[74] Tom Williams and Matthias Scheutz. 2019. Reference in Robotics: A Givenness Hierarchy Theoretic Approach. In *The Oxford Handbook of Reference*, Jeanette Gundel and Barbara Abbott (Eds.).

[75] Tom Williams, Daniel Szafir, and Tathagata Chakraborti. 2019. The Reality-Virtuality Interaction Cube. In *Proceedings of the 2nd International Workshop on Virtual, Augmented, and Mixed Reality for HRI*.

[76] Tom Williams, Daria Thames, Julia Novakoff, and Matthias Scheutz. 2018. "Thank You for Sharing that Interesting Fact!": Effects of Capability and Context on Indirect Speech Act Use in Task-Based Human-Robot Dialogue. In *Proceedings of the 13th ACM/IEEE International Conference on Human-Robot Interaction*.

[77] Tom Williams, Nhan Tran, Josh Rands, and Neil T. Dantam. 2018. Augmented, Mixed, and Virtual Reality Enabling of Robot Deixis. In *Proceedings of the 10th International Conference on Virtual, Augmented, and Mixed Reality*.