Towards an Understanding of Physical vs Virtual Robot Appendage Design

Zhao Han*, Albert Phan*, Amia Castro*, Fernando Sandoval Garza* and Tom Williams

MIRRORLab, Department of Computer Science, Colorado School of Mines, Golden, Colorado, USA 80401

zhaohan@mines.edu, aphan@mines.edu, acastro@mines.edu, fsandovalgarza@mines.edu, twilliams@mines.edu

Abstract—Augmented Reality (AR) or Mixed Reality (MR) enables innovative interactions by overlaying virtual imagery over the physical world. For roboticists, this creates new opportunities to apply proven non-verbal interaction patterns, like gesture, to physically-limited robots. However, a wealth of HRI research has demonstrated that there are real benefits to physical embodiment (compared, e.g., to virtual robots displayed on screens). This suggests that AR augmentation of virtual robot parts could lead to similar challenges.

In this work, we present the design of an experiment to objectively and subjectively compare the use of AR and physical arms for deictic gesture, in AR and physical task environments. Our future results will inform robot designers choosing between the use of physical and virtual arms, and provide new nuanced understanding of the use of mixed-reality technologies in HRI contexts.

Index Terms—augmented reality (AR), mixed reality (MR), deictic gesture, non-verbal communication, physical embodiment, robotics, mobile robots, human-robot interaction (HRI)

I. INTRODUCTION

To gain trust and acceptance, robots must be able to effectively communicate with people. Due to robots' unique physical embodiment [1], human-robot interaction (HRI) researchers have investigated non-verbal behaviors [2], such as implicit arm movement (e.g., [3], [4]), gestures [5], and eye gaze [6], [7]. Multimodal approaches pairing these nonverbal displays with verbal communication have also been well-studied (e.g., [8]–[10]). Results show that non-verbal behaviors themselves are particularly important as they increase task efficiency [6] and improve subjective perceptions of robots [3].

Unfortunately, most robot systems – such as mobile or telepresence robots, autonomous vehicles, or free-flying drones – do not have the physical morphology to express these types of nonverbal cues, lacking heads and eyes for gazing, or arms for gesturing. Moreover, the high degree-of-freedom requirements and complex mechanics of these morphological components, especially physical arms, present cost barriers, especially when such components would only be used for gesturing and not for manipulation. Finally, inclusion of physical components like arms presents well-known safety concerns [11].

To address these challenges, researchers have investigated *virtual* counterparts. For nonverbal facial cues, this has taken a variety of forms. The Furhat robot head [12], for example, uses projection mapping to display a humanlike face without
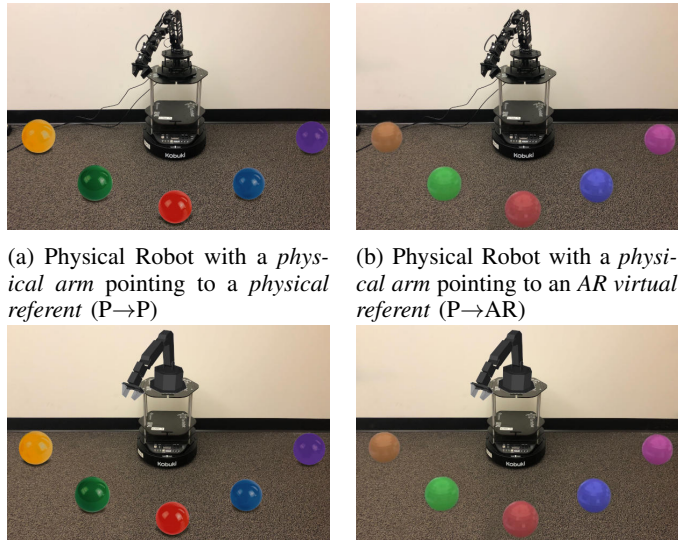
(a) Physical Robot with a *physical arm* pointing to a *physical referent* (P→P)

(b) Physical Robot with a *physical arm* pointing to an *AR virtual referent* (P→AR)

(c) Physical Robot with a *AR virtual arm* pointing to a *physical referent* (AR→P)

(d) Physical Robot with a *AR virtual arm* pointing to an *AR virtual referent* (AR→AR)

Fig. 1: Artist's rendering of the four conditions to be investigated. To investigate the intersection of physical and AR worlds, with a focus on referring behavior (physical/AR virtual *arm* × physical/AR virtual *referent*), we present an experiment design to evaluate such four interactions at the crossing of the real and mixed worlds. Note that the virtual model will be made hollow, like the physical real arm.

requiring actuation. Similarly, many approaches use tablets to display a robot's face (e.g., [13], [14]). Finally, some researchers have leveraged augmented reality to display and customize robots' faces [15].

Augmented Reality (AR) has also been recently used to provide a lower-cost alternative for enabling gestural capabilities. For example, Groechel et al. [11] studied the use of an AR arm on a mobile robot, and Hamilton et al. [16] further considered the use of AR arms for deictic gesture, comparing AR arms to other types of AR annotations (e.g., arrows [9]). Results showed that arms were better perceived subjectively. Yet, the performance differences between virtual (AR) and physical arms has not yet been explored. This means that while the monetary cost differences between these options can be readily compared, the performance differences between these platforms is not yet well understood, presenting a challenge for robot developers.

Moreover, it is further unclear whether differences between virtual and physical arms might be contingent on the virtuality or physicality of the task-relevant objects to which a robot might choose to gesture. It could be the case, for example, that virtual arms might be viewed more positively when used in tasks involving virtual referents, and vice versa. This would be a potentially complex challenge to reason over given that mixed-reality task environments may contain a mixture of virtual and physical objects.

In this paper, we present a study design to investigate the differences in objective performance and subjective perception between physical and virtual (AR) arms, as mediated by the physicality or virtuality of the robot's target referent (See Figure 1). This work will help robot designers to better understand whether and when to employ virtual rather than physical morphological components. Moreover, this will help provide more precise design guidelines that are sensitive to the nuances of mixed-reality robotics environments.

II. RELATED WORK

A. Virtual vs. Physical Agents

Much HRI research has already demonstrated differences in objective performance and subjective perception between purely virtual and purely physical robotic entities. Much of this work has compared physically embodied robots to virtual agents depicted on screens. This research demonstrates that embodied physical presence leads to greater influence [17], learning outcomes [18], task performance [19], [20], gaze following from infants [21], proximity [17], exercise [22], pervasiveness [23], positive perception [23], and social facilitation [24], forgiveness [24], enjoyableness [1], [22], [25], helpfulness [22], [26], and social attractiveness [22]. However, these works have not considered morphologies that blend the physical and the virtual, as is enabled by AR Technologies.

B. Virtual Agents in Augmented Reality

Different from the virtual agents wholly residing inside the virtual world, AR allows virtual objects/agents to be projected onto a user's view of the real world [27]. A variety of research has examined how interactants perceive agents in AR. Obaid et al. [28] showed that AR agent is perceived as physically distant where participants adjust their sound level accordingly, and Kim et al. [29] showed that AR agents aware of physical space were rated higher in social presence.

Other researchers have examined how people perceive virtual humans (ostensibly) interacting with the physical world through AR. Lee et al. [30] studied AR-visualized humans subtly moving a physical table in terms of presence, co-presence, and attentional allocation. Schmidt et al. [31] experimented with virtual humans manipulating physical objects: hitting a physical ball with a virtual golf club, but did not find statistical significance in realism and emotional responses. In contrast, our work considers a physical entity (a physical robot) with a virtual appendage, rather than a wholly virtual agent.

C. Augmented Reality for Robot Communication

In this work we are moreover specifically interested in robots using AR appendages for the purpose of communication. There has been a variety of work on AR for robot communication within the broader area of VAM-HRI [32], [33]. Frank et al. [34] used AR to show reachable spatial regions, to allow for humans to know where and when to pass objects to robots. Taylor et al. [35] used AR to remove robot arm occlusion and see occluded objects by making the arm transparent. Diehl et al. [36] used AR to verify learned behavior in the robot learning domain to increase safety and trust. In addition to AR using headsets, researchers have investigated projected AR. For example, [37] used projected AR to project car door frame, moving instructions, and task success in a car-assembly collaborative application. And Han et al. has focused on open science, making projector-based AR more readily available [38], [39].

D. Human and Robot Deictic Gesture

Finally, within this broader area, our work specifically examines robots' use of AR visualizations for the purposes of deictic gesture. Deictic gesture has been a topic of sustained and intense study both in human-human interaction [40], [41] and human-robot interaction [5]. Deictic gesture is a key and natural communication modality, with humans starting to use deictic gesture around 9-12 months [42], and mastering it around age 4 [43]. Adults continue to use deixis to direct interlocutor attention, so as to establish joint and shared attention [44]. As a non-verbal modality, gesturing is especially helpful in public noisy environments such as factories, warehouses, or malls [2], [45]. Accordingly, roboticists have leveraged this and study its effects for better understandable robots, e.g., in tabletop environments [46] and free-form direction-giving [47]. Specifically, research shows that robot, like humans, can shift interlocutor attention [48] and can use a variety of deictic gestures, not only pointing [5], [49]. Williams et al. have begun to explore the use of deictic gesture within Augmented Reality [9], [50]–[53], although most of this work has been with non-anthroporphic visualizations like virtual arrows. In contrast, Hamilton et al. [16], like ourselves in this work, recently examined virtual arms, and showed that AR virtual arm enhanced social presence, likability. Unlike Hamilton, however, we are interested in explicitly comparing virtual arms to physical arms (rather than other types of virtual gestures) and in understanding the role that the physical or virtual nature of the environment might mediate these differences.

III. HYPOTHESES

We approach this work with a set of key hypotheses and expectations. First, we maintain hypotheses regarding the objective effectiveness of AR gestures.

Hypothesis 1 (H1) – Equal Accuracy. We believe that robot deixis with a virtual arm will be no less accurate than deixis with a physical arm.

Hypothesis 2 (H2) – Reality Alignment mediates Efficiency. We believe that while using physical or virtual arms to

refer to physical or virtual referents, *respectively*, should have equivalent efficiency, we hypothesize that a mismatch between these levels of reality (i.e., virtual arms pointing to physical objects, and vice versa) could increase the time needed by users to identify the robot's target, due to a need for additional cognitive processing to explicitly overcome this misalignment.

Hypothesis 3 (H3) – Reality Mediates Perception. Similarly, we believe a mismatch between gesture-reality and referent-reality will reduce perceived naturality and likability.

IV. METHOD

To investigate these hypotheses, we plan to employ the following experimental method.

A. Apparatus

1) Robot Platform: As one of the application domains of this work is a gesturally-limited robot, we plan to use a TurtleBot 2 mobile robot [54]. This differential wheeled robots is the second generation of the Turbot family, and is maintained by the current maintainer of Robot Operating System (ROS) [55], thus having a large support community. The specification for TurtleBot compatible platforms can be found on ros.org [56].

2) Mixed-Reality Head-Mounted Display (MR-HMD): Microsoft HoloLens 2 [57] will be used as our MR-HMD. It is a commercial-grade see-through holographic mixed reality headset with a $43° × 29°$ Field of View (FOV).

3) Physical/Virtual Robot Arm: The physical arm we will use will be the WidowX Robot Arm [58]: a 5-DoF arm with a parallel gripper, which an reach up to $41cm$ horizontally and $44cm$ vertically. Our virtual arm will be created using the CAD models and Unified Robot Description Format (URDF) model of this arm [59], as rendered in Unity. Although the WidowX is a relatively simple arm[1], it is relatively costly as a deictic appendage, priced at $1,699.95 USD.

The virtual arm will have the same distance to the TurtleBot top when rendered in Unity. To affix the AR virtual arm to the same position as the physical arm, we plan to place a $12cm$ cardboard cube on the top panel of the TurtleBot2. Each face of the cube has fiducial markers for localization [60].

4) Physical/Virtual Referents: Five spheres [61] will be used as communicative referents and will be arced within the field of view of HoloLens 2 (See Figure 1). Each sphere measures $d = 15.24cm$ ($6in$) in diameter and will be placed $45°$ apart, as shown in Fig 1. The distances between the robot and the referents are preliminarily determined to be $3 × d$ for clarity and $1m$ within field of view. The physical and rendered spheres will have the same size and placement.

B. Gesturing Task and Implementation

These materials will be used in the context of a standard gesture-comprehension experiment. In each trial, the WidowX robot arm, mounted or simulated on top of TurtleBot 2, will randomly point to one of the colored spherical targets, which

---

[1]See other arms on https://www.trossenrobotics.com/robotic-arms.aspx for a comparison of prices.

TABLE I: Four Study Conditions across Two Dimensions on Physicality/Virtuality

		Referent Virtuality	
		Physical	*AR Virtual*
Arm Virtuality	*Physical*	P→P	P→AR
	AR Virtual	AR→P	AR→AR

participants will then be asked to identify by air-tapping on that target. This will be repeated ten times, with targets chosen at random. While a controller could be used, the four directional buttons do not work well for five referents, and would introduce confounds for measuring response time.

For each gesture, the MoveIt motion planning framework [62] will be used to move the end effector to the desired pointing pose. As we plan to conduct this experiment in person, the trajectory generated by MoveIt to its final pose, non-deterministic due to the probabilistic algorithms [63], will be made deterministic by specifying multiple waypoints before the pointing poses towards different spheres. This approach to a deterministic outcome has seen success in robot-to-human handover tasks [14]. As an alternative approach, we are also investigating recording the trajectory output by MoveIt given both the placement of the robot and the spheres are static, thus the recorded trajectory will be replayed at experiment time.

For the AR virtual arm, i.e., the WidowX arm model rendered in Unity, we plan to implement the functionality that converts the MoveIt trajectory to key-frame-based animation either by subscribing to MoveIt output trajectory directly or from a pre-recorded trajectory.

The MoveIt and Unity implementation will be open-sourced on GitHub to facilitate replication.

C. Experiment Design

This study will follow a $2 × 2$ *between-subjects* design because a within-subjects design would require rapidly uninstalling the physical arm and cardboard cube from the top panel of the robot, which would take significant effort and time, and could be error-prone within the short timeframe of an experiment.

As implied throughout this work so far, we will manipulate whether the arm and the referent, i.e., the spheres, are physical or rendered. Formally, two independent variables will be manipulated: *referee virtuality* and *referent virtuality*. Thus, there will be four study conditions across the two factors:

- P→P (Physical): *Real* arm pointing at *real* spheres
- P→AR: *Real* arm pointing at *virtual* spheres
- AR→P': *Virtual* AR arm pointing at *real* spheres
- AR→AR: *Virtual* AR arm pointing at *virtual* spheres

D. Procedure

The study will be conducted in person in order to use the head-mounted display. All apparatus will be disinfected before use due to COVID-19 concerns.

Upon arrival, participants will be presented with informed consent information, in which they will be asked to identify the object the robot is referring to quickly and accurately to

impose the same amount of time pressure. After agreeing to participate, they will fill out a demographic survey and be randomly assigned to one of the four experimental conditions.

After watching a video on how to put on HoloLens 2, they will wear the headset and receive further training. The experimenter will first briefly get participants familiar with the task again as described in the informed consent form. Then participants will experience an interactive tutorial on HoloLens 2, designed in Unity. It will 1) allow participants walk through sample experiment trials to see how either the physical or virtual arm moves to gesture, 2) allow eye-tracking user calibration to accurately collect accuracy data (See Section IV-E1 below), and 3) get participants familiar with the air-tap gesturing to confirm the target object they believe in. While this is an onboarding experience, it is also considered to avoid novel effects. Experimenters will ask clarifying questions to ensure their understanding of the task and the procedure. After completing 10 trials, they will be asked to answer a questionnaire with all the subjective measures. At the end, participants will be paid according to our planned pilot study duration and debriefed.

E. Data Collection and Measures

To test our hypotheses, we plan to collect two objective metrics and five subjective metrics to capture experience. Some of the subjective metrics are inspired by [16].

1) Accuracy: To facilitate data gathering for accuracy to measure effectiveness, participants in every condition will wear an HoloLens 2 in order to assess whether they have inferred which referent the robot has pointed at. This is achieved by the built-in eye-tracking API [64] of HoloLens 2. To collect such gazing information, the same numbers of invisible objects as the referents' will be implemented in Unity to be at the same positions as the rendered balls.

Accuracy will be calculated as the percentage of true positives where participants looked at the target referent in all trials, confirmed by "clicking" (using air-tap gesture) on the referent in their belief. The gesture will be detected by HoloLens 2's built-in gesture recognition capabilities.

2) Reaction Time: Similar to how accuracy is measured, reaction time will be measured through the HoloLens 2 eye-tracking API. Specifically, reaction time will be the duration between when the robot arm starts moving from its home position to when participants look at the target object.

3) Social Presence: As seen in Section II, social presence, defined as the feeling of being in the company of another social actor [65], has been a central metric in studies involving virtual agents. It can enable more effective social and group interactions [66], [67]. Within HRI, it has been found to increase enjoyment and desire to re-interact [68].

4) Anthropomorphism: Projecting human characteristics to non-human entities [69]–[71], such as attaching the AR virtual arm to the TurtleBot 2 in this work, encourage humans to re-use the familiar interaction patterns from human-human interactions. It facilitates sensemaking and mental model alignment [71], leading humans to be more willing to interact, accept, and understand robot's behaviors [72]. Robots that use gesture have been found to appear more anthropomorphic [73]. Hamilton et al. [16] suggested that the robot with a virtual arm may have been viewed as more anthropomorphic, however it is unclear how this is compared with a robot with a physical arm. Anthropomorphism will be measured using the Godspeed Anthropomorphism scale [74].

5) Likability: As one of the primary metrics used in nonverbal robot communication [51], [73], [75], Likability summarize peoples' overall perceptions of technology, key to estimate people's experience. Similarly, Hamilton et al. [16] found evidence that the robot with a virtual arm enhanced likability, but it did not compare with the physical counterpart. Likability will be measured using the Godspeed Likability scale [74].

6) Warmth and Competence: As psychological constructs at the core of social judgment, warmth and competence are responsible for social perceptions among humans [76]. Warmth captures whether an actor is sociable and well-intentioned, and competence captures whether they can deliver on those intentions. Warmth and competence are thus key predictors of effective and preferable interactions, both for human-human interaction [76] and human-robot interaction [77], [78]. Moreover, they have been connected to social presence [79], and anthropomorphism [80], [81]. Warmth and Competence will be measured using the ROSAS Scale [82].

F. Data Analysis and Participants

We plan to analyze the data within a Bayesian analysis framework [83] using the JASP 1.6 (version at submission time, will update) software package [84], with the default settings justified by Wagenmakers et al. [85]. Bayesian analysis with Bayes factors has benefits over the more common frequentist approach [83]. Some of them include that the Bayes factors can quantify evidence *for* the null hypothesis $\mathcal{H}_0$, and evidence *for* $\mathcal{H}_1$ vs. $\mathcal{H}_0$. In contrast, the p value cannot provide a measure of evidence in favor of $\mathcal{H}_0$. For more details, we refer readers to [83]. All experimental data and analysis scripts will be made available publicly to facilitate replication.

Unlike the frequentist approach in need of power analysis to achieve sufficient power [86], [87], the Bayesian analysis does not strictly require power analyses to determine sample size [88] as it is not dependent on the central limit theorem. Nonetheless, we plan to recruit at least 25 participants, similar to the AR virtual arm vs. the AR virtual arrow study [16].

V. CONCLUSION

In this workshop paper, we have described the design of an experiment designed to investigate the differences in performance between physically-limited robots that use either physical or virtual (AR) arms, as mediated by the physical or virtual nature of their target referents. Our immediate future work will be to conduct, analyze, and report the results of this experiment. Our hope is that our results provide new insights into mixed-reality human-robot interaction, and will help inform robot designers' decisions as to whether to use virtual or physical robotic arms.

REFERENCES

[1] J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric, "The role of physical embodiment in human-robot interaction," in *The 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2006, pp. 117–122.

[2] E. Cha, Y. Kim, T. Fong, M. J. Mataric *et al.*, "A survey of nonverbal signaling methods for non-humanoid robots," *Foundations and Trends® in Robotics*, vol. 6, no. 4, pp. 211–323, 2018.

[3] A. D. Dragan, S. Bauman, J. Forlizzi, and S. S. Srinivasa, "Effects of robot motion on human-robot collaboration," in *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2015, pp. 51–58.

[4] M. Kwon, S. H. Huang, and A. D. Dragan, "Expressing robot incapability," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 87–95.

[5] A. Sauppé and B. Mutlu, "Robot deictics: How gesture and context shape referential communication," in *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2014, pp. 342–349.

[6] A. Moon, D. M. Troniak, B. Gleeson, M. K. Pan, M. Zheng, B. A. Blumer, K. MacLean, and E. A. Croft, "Meet me where i'm gazing: how shared attention gaze affects human-robot handover timing," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 2014, pp. 334–341.

[7] H. Admoni and B. Scassellati, "Social eye gaze in human-robot interaction: a review," *Journal of Human-Robot Interaction*, vol. 6, no. 1, pp. 25–63, 2017.

[8] P. Bremner and U. Leonards, "Iconic gestures for robot avatars, recognition and integration with speech," *Frontiers in psychology*, vol. 7, p. 183, 2016.

[9] T. Williams, M. Bussing, S. Cabrol, E. Boyle, and N. Tran, "Mixed reality deictic gesture for multi-modal robot communication," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2019, pp. 191–201.

[10] Z. Han and H. A. Yanco, "Explaining a robot's past: Communicating missing causal information by replay, speech and projection," *ACM Transactions on Human-Robot Interaction (THRI)*, Under review as of Jan. 2022.

[11] T. Groechel, Z. Shi, R. Pakkar, and M. J. Matarić, "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, 2019, pp. 1–8.

[12] S. Al Moubayed, J. Beskow, G. Skantze, and B. Granström, "Furhat: a back-projected human-like robot head for multiparty human-machine interaction," in *Cognitive behavioural systems*. Springer, 2012, pp. 114–130.

[13] N. T. Fitter and K. J. Kuchenbecker, "Designing and assessing expressive open-source faces for the baxter robot," in *International Conference on Social Robotics*. Springer, 2016, pp. 340–350.

[14] Z. Han and H. Yanco, "The effects of proactive release behaviors during human-robot handovers," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2019, pp. 440–448.

[15] S. M. zu Borgsen, P. Renner, F. Lier, T. Pfeiffer, and S. Wachsmuth, "Improving human-robot handover research by mixed reality techniques," *VAM-HRI*, pp. 2018–03, 2018.

[16] J. Hamilton, T. Phung, N. Tran, and T. Williams, "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*, 2021, pp. 177–186. [Online]. Available: https://mirrorlab.mines.edu/wp-content/uploads/sites/198/2020/12/hamilton2021hri.pdf

[17] W. A. Bainbridge, J. W. Hart, E. S. Kim, and B. Scassellati, "The benefits of interactions with physically present robots over video-displayed agents," *International Journal of Social Robotics*, vol. 3, no. 1, pp. 41–52, 2011.

[18] J. Kennedy, P. Baxter, and T. Belpaeme, "Children comply with a robot's indirect requests," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 2014, pp. 198–199.

[19] J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric, "Embodiment and human-robot interaction: A task-based perspective," in *The 16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2007, pp. 872–877.

[20] D. Leyzberg, S. Spaulding, M. Toneva, and B. Scassellati, "The physical presence of a robot tutor increases cognitive learning gains," in *Proceedings of the annual meeting of the cognitive science society*, vol. 34, no. 34, 2012.

[21] A. N. Meltzoff, R. Brooks, A. P. Shon, and R. P. Rao, ""social" robots are psychological agents for infants: A test of gaze following," *Neural networks*, vol. 23, no. 8-9, pp. 966–972, 2010.

[22] J. Fasola and M. J. Matarić, "A socially assistive robot exercise coach for the elderly," *Journal of Human-Robot Interaction*, vol. 2, no. 2, pp. 3–32, 2013.

[23] J. Li, "The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents," *International Journal of Human-Computer Studies*, vol. 77, pp. 23–37, 2015.

[24] C. Bartneck, "Interacting with an embodied emotional character," in *Proceedings of the 2003 international conference on Designing pleasurable products and interfaces*, 2003, pp. 55–60.

[25] K. Pollmann, C. Ruff, K. Vetter, and G. Zimmermann, "Robot vs. voice assistant: Is playing with pepper more fun than playing with alexa?" in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 2020, pp. 395–397.

[26] A. Powers, S. , S. Fussell, and C. Torrey, "Comparing a computer agent with a humanoid robot," in *Proceedings of the 2007 ACM/IEEE international conference on Human-robot interaction*, 2007, pp. 145–152.

[27] I. Wang, J. Smith, and J. Ruiz, "Exploring virtual agents for augmented reality," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[28] M. Obaid, R. Niewiadomski, and C. Pelachaud, "Perception of spatial relations and of coexistence with virtual agents," in *International Workshop on Intelligent Virtual Agents*. Springer, 2011, pp. 363–369.

[29] K. Kim, G. Bruder, and G. Welch, "Exploring the effects of observed physicality conflicts on real-virtual human interaction in augmented reality," in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, 2017, pp. 1–7.

[30] M. Lee, K. Kim, S. Daher, A. Raij, R. Schubert, J. Bailenson, and G. Welch, "The wobbly table: Increased social presence via subtle incidental movement of a real-virtual table," in *2016 IEEE Virtual Reality (VR)*. IEEE, 2016, pp. 11–17.

[31] S. Schmidt, O. J. A. Nunez, and F. Steinicke, "Blended agents: Manipulation of physical objects within mixed reality environments and beyond," in *Symposium on Spatial User Interaction*, 2019, pp. 1–10.

[32] M. Walker, T. Phung, T. Chakraborti, T. Williams, and D. Szafir, "Virtual, augmented, and mixed reality for human-robot interaction: A survey and virtual design element taxonomy," *arXiv preprint arXiv:2202.11249*, 2022.

[33] T. Williams, D. Szafir, T. Chakraborti, and H. Ben Amor, "Virtual, augmented, and mixed reality for human-robot interaction," in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 403–404.

[34] J. A. Frank, M. Moorhead, and V. Kapila, "Mobile mixed-reality interfaces that enhance human–robot interaction in shared spaces," *Frontiers in Robotics and AI*, vol. 4, p. 20, 2017.

[35] A. V. Taylor, A. Matsumoto, E. J. Carter, A. Plopski, and H. Admoni, "Diminished reality for close quarters robotic telemanipulation," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 11 531–11 538.

[36] M. Diehl, A. Plopski, H. Kato, and K. Ramirez-Amaro, "Augmented reality interface to verify robot learning," in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2020, pp. 378–383.

[37] R. K. Ganesan, Y. K. Rathore, H. M. Ross, and H. B. Amor, "Better teaming through visual cues: how projecting imagery in a workspace can improve human-robot collaboration," *IEEE Robotics & Automation Magazine*, vol. 25, no. 2, pp. 59–71, 2018.

[38] Z. Han, A. Wilkinson, J. Parrillo, J. Allspaw, and H. A. Yanco, "Projection mapping implementation: Enabling direct externalization of perception results and action intent to improve robot explainability," in *2020 AAAI Fall Symposium on The Artificial Intelligence for Human-Robot Interaction (AI-HRI)*, 2020.

[39] Z. Han, J. Parrillo, A. Wilkinson, H. A. Yanco, and T. Williams, "Projecting robot navigation paths: Hardware and software for projected ar," in *2022 ACM/IEEE International Conference on Human-Robot Interaction (HRI), Short Contributions*, 2022.

[40] S. C. Levinson, "5 deixis," *The handbook of pragmatics*, p. 97, 2004.

[41] S. Norris, "Three hierarchical positions of deictic gesture in relation to spoken language: a multimodal interaction analysis," *Visual Communication*, vol. 10, no. 2, pp. 129–147, 2011.

[42] E. Bates, *Language and context: The acquisition of pragmatics*. Academic Press, 1976.

[43] E. V. Clark and C. Sengul, "Strategies in the acquisition of deixis," *Journal of child language*, vol. 5, no. 3, pp. 457–475, 1978.

[44] A. Bangerter and M. M. Louwerse, "Focusing attention with deictic gestures and linguistic expressions," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 27, no. 27, 2005.

[45] S. Harrison, "The creation and implementation of a gesture code for factory communication," in *GESPIN 2011: Gesture and speech in interaction*, 2011.

[46] M. Salem, S. Kopp, I. Wachsmuth, K. Rohlfing, and F. Joublin, "Generation and evaluation of communicative robot gesture," *International Journal of Social Robotics*, vol. 4, no. 2, pp. 201–217, 2012.

[47] Y. Okuno, T. Kanda, M. Imai, H. Ishiguro, and N. Hagita, "Providing route directions: design of robot's utterance, gesture, and timing," in *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2009, pp. 53–60.

[48] A. G. Brooks and C. Breazeal, "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*, 2006, pp. 297–304.

[49] H. H. Clark, "Coordinating with each other in a material world," *Discourse studies*, vol. 7, no. 4-5, pp. 507–525, 2005.

[50] T. Williams, N. Tran, J. Rands, and N. T. Dantam, "Augmented, mixed, and virtual reality enabling of robot deixis," in *Proceedings of the 10th International Conference on Virtual, Augmented, and Mixed Reality*, 2018.

[51] T. Williams, M. Bussing, S. Cabrol, I. Lau, E. Boyle, and N. Tran, "Investigating the potential effectiveness of allocentric mixed reality deictic gesture," in *Proceedings of the 11th International Conference on Virtual, Augmented, and Mixed Reality*, 2019.

[52] T. Williams, L. Hirshfield, N. Tran, T. Grant, and N. Woodward, "Using augmented reality to better study human-robot interaction," in *International Conference on Human-Computer Interaction*. Springer, 2020, pp. 643–654.

[53] N. Tran, T. Grant, T. Phung, L. Hirshfield, C. Wickens, and T. Williams, "Get this?↓ mixed reality improves robot communication regardless of mental workload," in *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 2021, pp. 412–416.

[54] "TurtleBot2: Open-source robot development kit for apps on wheels," https://www.turtlebot.com/turtlebot2/, online; accessed 2022-02-12.

[55] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Ng *et al.*, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, no. 3.2. Kobe, Japan, 2009, p. 5.

[56] "Specification for TurtleBot Compatible Platforms," https://www.ros.org/reps/rep-0119.html, online; accessed 2022-02-12.

[57] "About HoloLens 2," https://docs.microsoft.com/en-us/hololens/hololens2-hardware, online; accessed 2022-02-15.

[58] "WidowX Robot Arm Kit," https://www.trossenrobotics.com/widowxrobotarm/, online; accessed 2022-02-12.

[59] "The widowx arm description ROS package," https://github.com/AnasIbrahim/widowx_arm/tree/master/widowx_arm_description, online; accessed 2022-02-15.

[60] E. Olson, "Apriltag: A robust and flexible visual fiducial system," in *2011 IEEE international conference on robotics and automation*. IEEE, 2011, pp. 3400–3407.

[61] "Lakeshore 6" Activity Balls," https://www.lakeshorelearning.com/products/ca/p/RA298/, online; accessed 2022-02-15.

[62] S. Chitta, I. Sucan, and S. Cousins, "Moveit![ros topics]," *IEEE Robotics & Automation Magazine*, vol. 19, no. 1, pp. 18–19, 2012.

[63] I. A. Sucan, M. Moll, and L. E. Kavraki, "The open motion planning library," *IEEE Robotics & Automation Magazine*, vol. 19, no. 4, pp. 72–82, 2012.

[64] "Eye tracking on HoloLens 2," https://docs.microsoft.com/en-us/windows/mixed-reality/design/eye-tracking, online; accessed 2022-02-15.

[65] P. Skalski and R. Tamborini, "The role of social presence in interactive agent-based persuasion," *Media psychology*, vol. 10, no. 3, pp. 385–413, 2007.

[66] F. Biocca, C. Harms, and J. K. Burgoon, "Toward a more robust theory and measure of social presence: Review and suggested criteria," *Presence: Teleoperators & virtual environments*, vol. 12, no. 5, pp. 456–480, 2003.

[67] M. Lombard and T. Ditton, "At the heart of it all: The concept of presence," *Journal of computer-mediated communication*, vol. 3, no. 2, p. JCMC321, 1997.

[68] M. Heerink, B. Kröse, V. Evers, and B. Wielinga, "Assessing acceptance of assistive social agent technology by older adults: the almere model," *International journal of social robotics*, vol. 2, no. 4, pp. 361–375, 2010.

[69] B. R. Duffy, "Anthropomorphism and robotics," *The society for the study of artificial intelligence and the simulation of behaviour*, vol. 20, 2002.

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

[71] C. DiSalvo and F. Gemperle, "From seduction to fulfillment: the use of anthropomorphic form in design," in *Proceedings of the 2003 international conference on Designing pleasurable products and interfaces*, 2003, pp. 67–72.

[72] D. Kuchenbrandt, N. Riether, and F. Eyssel, "Does anthropomorphism reduce stress in hri?" in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 2014, pp. 218–219.

[73] M. Salem, F. Eyssel, K. Rohlfing, S. Kopp, and F. Joublin, "To err is human (-like): Effects of robot gesture on perceived anthropomorphism and likability," *International Journal of Social Robotics*, vol. 5, no. 3, pp. 313–323, 2013.

[74] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International journal of social robotics*, vol. 1, no. 1, pp. 71–81, 2009.

[75] T. Williams, D. Thames, J. Novakoff, and M. Scheutz, ""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*, 2018.

[76] S. T. Fiske, A. J. Cuddy, and P. Glick, "Universal dimensions of social cognition: Warmth and competence," *Trends in cognitive sciences*, vol. 11, no. 2, pp. 77–83, 2007.

[77] M. M. Scheunemann, R. H. Cuijpers, and C. Salge, "Warmth and competence to predict human preference of robot behavior in physical human-robot interaction," *arXiv preprint arXiv:2008.05799*, 2020.

[78] C. M. Carpinella, A. B. Wyman, M. A. Perez, and S. J. Stroessner, "The robotic social attributes scale (rosas) development and validation," in *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*, 2017, pp. 254–262.

[79] K. Hassanein and M. Head, "Manipulating perceived social presence through the web interface and its impact on attitude towards online shopping," *International Journal of Human-Computer Studies*, vol. 65, no. 8, pp. 689–708, 2007.

[80] S. Y. Kim, B. H. Schmitt, and N. M. Thalmann, "Eliza in the uncanny valley: anthropomorphizing consumer robots increases their perceived warmth but decreases liking," *Marketing letters*, vol. 30, no. 1, pp. 1–12, 2019.

[81] A. Waytz, J. Heafner, and N. Epley, "The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle," *Journal of Experimental Social Psychology*, vol. 52, pp. 113–117, 2014.

[82] C. M. Carpinella, A. B. Wyman, M. A. Perez, and S. J. Stroessner, "The robotic social attributes scale (rosas) development and validation," in *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*, 2017, pp. 254–262.

[83] E.-J. Wagenmakers, M. Marsman, T. Jamil, A. Ly, J. Verhagen, J. Love, R. Selker, Q. F. Gronau, M. Šmíra, S. Epskamp *et al.*, "Bayesian inference for psychology. part i: Theoretical advantages and practical ramifications," *Psychonomic bulletin & review*, vol. 25, no. 1, pp. 35–57, 2018.

[84] JASP Team, "JASP (Version 0.16)[Computer software]," 2021. [Online]. Available: https://jasp-stats.org/

[85] E.-J. Wagenmakers, J. Love, M. Marsman, T. Jamil, A. Ly, J. Verhagen, R. Selker, Q. F. Gronau, D. Dropmann, B. Boutin *et al.*, "Bayesian inference for psychology. part ii: Example applications with jasp," *Psychonomic bulletin & review*, vol. 25, no. 1, pp. 58–76, 2018.

[86] K. S. Button, J. Ioannidis, C. Mokrysz, B. A. Nosek, J. Flint, E. S. Robinson, and M. R. Munafò, "Power failure: why small sample size undermines the reliability of neuroscience," *Nature reviews neuroscience*, vol. 14, no. 5, pp. 365–376, 2013.

[87] M. E. Bartlett, C. Edmunds, T. Belpaeme, and S. Thill, "Have i got the power? analysing and reporting statistical power in hri," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 11, no. 2, pp. 1–16, 2022.

[88] J. Correll, C. Mellinger, G. H. McClelland, and C. M. Judd, "Avoid cohen's 'small','medium', and 'large'for power analysis," *Trends in Cognitive Sciences*, vol. 24, no. 3, pp. 200–207, 2020.