

# Exploring Familiar Design Strategies to Explain Robot Vision Capabilities

Hong Wang  
RARE Lab

University of South Florida  
Tampa, Florida, USA  
hongw@usf.edu

Maria Julia Vidal  
RARE Lab

University of South Florida  
Tampa, Florida, USA  
mariajulia57@usf.edu

Zhao Han  
RARE Lab

University of South Florida  
Tampa, Florida, USA  
zhaohan@usf.edu

**Abstract**—Humans often assume that robots share the same field of view (FoV) as themselves, given their human-like appearance. In reality, robots have a much narrower FoV (e.g., Pepper robot’s 54.4° and Fetch robot’s 54°) than humans’ 180°, leading to misaligned mental models and reduced efficiency in collaborative tasks. For instance, a user might place an object just outside the robot’s narrow FoV, assuming it can see but only to find the robot failing to retrieve it, resulting in confusion and disappointment, thus hindering human-robot interaction (HRI). To address this gap, we propose five indicators to proactively communicate the robot’s FoV. Rooted in familiar human experiences, the proposed designs span egocentric (robot-oriented) to allocentric (task-oriented) spaces: *Glasses* in the eye space are designed to physically block areas outside FoV, while *gestures* from head space (raise hands to eye sides) to task space (extend arms to workspace) dynamically indicate the vision range, with the head-mounted projector *spotlights* the FoV directly at the workspace. We will compare these designs with three baseline conditions in common modalities: No indicator, image on screen, and voice announcement. Through a kitting task, we will gather objective metrics for task performance (error rate, reaction time) and subjective perceptions to gauge user experience (likability, trust). Our future results will guide designers in integrating those designs for clear FoV communication, fostering efficient and trustworthy human-robot teams.

**Index Terms**—System transparency, robot explainability, robot design, vision capability, field of view (FoV), human-robot interaction (HRI)

## I. INTRODUCTION

Imagine a busy office where a humanoid robot is assisting employees by fetching supplies. A colleague places a box just beyond the robot’s line of sight, expecting it to detect and deliver the box on its own. When the robot fails to detect the object, confusion arises: Why cannot it see what is obviously visible to humans?

This disconnect is rooted in a fundamental misunderstanding: Because robots look human-like, people often assume that they have human-like vision capabilities. However, while humans have a horizontal field of view of more than 180°, a robot’s eyes or cameras typically have less than 60° on its horizontal FoV (e.g., Pepper’s 54.4° [1], NAO’s 56.3° [2] and Fetch’s 54° [3, 4]). Such a mismatch between human expectations and robot performance leads to inaccurate mental models [5] of robot vision capabilities, likely causing inefficiencies in

collaborative tasks, e.g., multiple research has shown repeated clarifications for objects during handovers [6, 7].

To address such misaligned expectations, we argue that robots should proactively explain their vision limitations not only through post-hoc verbal clarifications, but also via design-driven communication. This should be done during interactions or before interaction starts in order to gain trust, a critical element for fluent human-robot collaboration [8, 9]. For example, studies show that people are more likely to trust robots when they have a detailed understanding of robot capabilities [10, 11].

Furthermore, using familiar interaction patterns from human-human interactions encourages the alignment of mental models between human and robot, leading to improved acceptance and perceived trustworthiness [12]. Robots using gestures, for example, have been found to be perceived as having higher anthropomorphism than no gestures at all [13, 14].

Yet, trust not only depends on anthropomorphic traits, but also on clear communication of robot capabilities. Communication strategies in HRI are classified as implicit (e.g., gestures or movements) and explicit (e.g., verbal explanations) [15]. In practice, robots can provide verbal communication, like explanations, or non-verbal ones, like physically scanning their surroundings to accommodate limited vision. However, these strategies often require back-and-forth clarifications with human collaborators. Moreover, they fail to overcome the wrong yet long-lasting first assumptions that users make based on the robot’s appearance, e.g., wide FoV from a robot’s human-like eyes.

Thus, we aim to address such misalignment through robot design communication, activating familiar interaction patterns. They may be particularly helpful for robot users who will make the first impression on the robot or as a quick-start guide after robot deployment, where the robot’s capabilities can be communicated through familiar accessories, arm gestures, or other objects. By aligning humans’ expectations with the robot’s actual abilities, we hope these designs allow users to provide informed requests and reduce the need for repeated clarifications. We aim not only to correct mental model misalignment but also to maintain positive perceptions of the robot, such as its likability, competence, and trustworthiness.

Specifically, to communicate the robot’s FoV, we propose

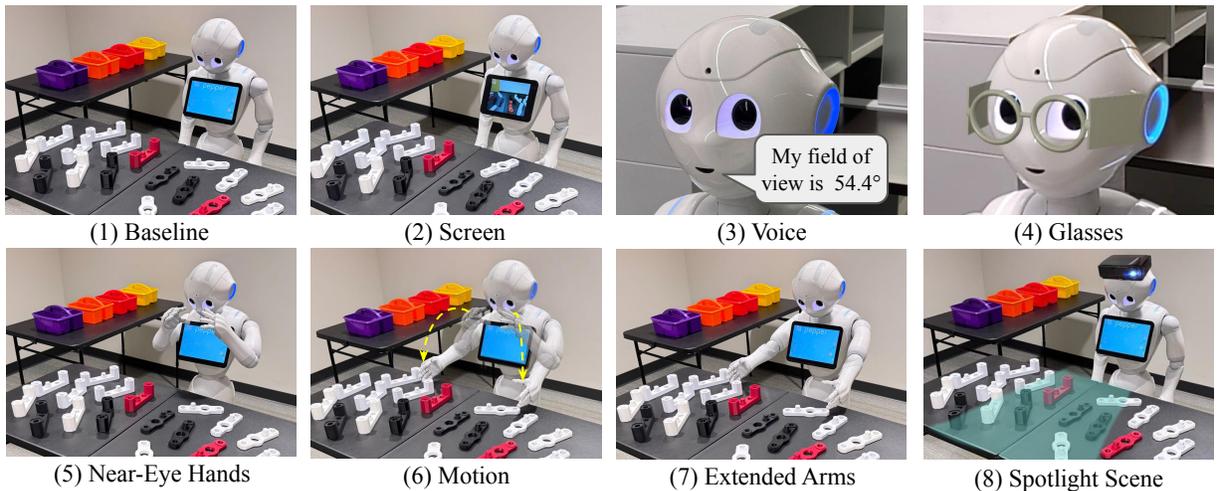


Fig. 1: The proposed designs to communicate the robot’s FoV with a baseline at the beginning, to be implemented to the real robot and evaluated in a human-subjects study. Designs are detailed in Section III.

five designs (Figure 1, 4–8) and plan to compare them with three baseline conditions in common modalities: No indicator, image on a screen, and voice (Figure 1, 1–3). Our designs draw inspiration from everyday experiences (e.g., wearing glasses, stage spotlights) and relatable behaviors (e.g., human gestures). We also present a human-subjects study design to evaluate them in a collaborative kitting scenario, measuring objective performance (e.g., error rates, reaction times) and subjective perceptions (e.g., trust, likability) to determine which designs best align user expectations with the robot’s actual vision capabilities and the trade-offs that robot designers may need to make.

## II. RELATED WORK

### A. Anthropomorphism and Familiar Experiences in HRI

Anthropomorphism is the human tendency to transfer human-like characteristics to non-human entities [16, 17, 18]. Research shows that anthropomorphic design in embodied robots can enhance subjective outcomes, such as likability, familiarity, and perceived competence, thereby improving interaction quality [19]. Physical embodiment, as opposed to virtual agents, often produces stronger positive effects on perception and performance during interactions [20, 21].

However, designing anthropomorphic robots requires careful consideration to balance its benefits with potential drawbacks. Many studies have explored the impact of anthropomorphic features through robot design. For instance, Fong et al. [22] provided an extensive review of anthropomorphic robot design, emphasizing that appropriate anthropomorphic cues such as facial expressions, gestures, and human-like bodily movements can improve user acceptance and engagement. Ahmad et al. [23] compared a human-like (Pepper) and a machine-like (Husky) robot in a game scenario, finding participants interacting with the humanoid robot reported lower cognitive load and higher trust. While anthropomorphism can support users’ perception of predictable robot behavior,

excessive anthropomorphic features may result in discomfort or reduced trust [24]. For example, Onnasch and Hildebrandt [25] found anthropomorphic features like eyes and eyebrows [25] found anthropomorphic features like eyes and eyebrows capture attention, but also distract people from relevant task, failed to reveal positive effects on trust in industrial HRI.

With most of the prior work focus on enhancing subjective outcomes, such as likability and engagement, few of them address how to effectively communicate a robot’s capabilities through anthropomorphic designs. Addressing this gap, our approach leverages anthropomorphic elements and familiar experience in designs, i.e., glasses (design 4) and spotlight scene (design 8), to help people connect to everyday life, making it easier to understand robot vision capabilities.

### B. Body Language and Gestures

While anthropomorphism is more focused on appearance [26], body language and human-like gestures are a communication style to design a robot’s behavior anthropomorphically [22, 25]. Robots that use gestures have been found to appear more anthropomorphic [13]. Specifically, gestures are a form of nonverbal communication that includes visible movements of the hands, arms, face, and other parts of the body to express an idea or meaning. The most straightforward way for robots to communicate effectively is to use similar nonverbal cues as humans (e.g., arm gesture) [27], which helps reduce cognitive effort and improve trust. For example, Riek et al. [28] found that cooperative gestures aligned with human expectations can reduce discomfort and enhance interaction effectiveness. Similarly, Xu et al. [29] used body language to express robot NAO’s mood in an imitation game. They found people were able to differentiate between positive and negative robot mood. Sauppé and Mutlu [30] showed that human-like deictic gestures, such as pointing, touching, and presenting on robots enable users to better understand references in complex settings. Dragan et al. [31] found that motion designed to



Fig. 2: Spectrum of indicator designs ranging from egocentric (robot) to allocentric (task), and vice versa.

convey the robot’s intent clearly leads to more fluent collaborations and higher comfort compared to functional and predictable motion. In addition, Block et al. [32] found that intra-hug gestures, such as holding, rubbing, patting, and squeezing, improved user experience by responding to user actions and initiating gestures proactively. This mirroring of human hugging dynamics made interactions more natural and emotionally engaging. Besides, augmented reality and abstract pointing behaviors have also been studied [21, 33].

These works highlight the importance of integrating human-like behaviors in robot design to enhance user understanding and trust. Building on these findings, our work explores both static and dynamic gesture-based strategies (designs 5-7) to communicate the robot’s narrower FoV, bridging the gap between its human-like appearance and actual vision capabilities.

### III. TAXONOMY AND DESIGNS

To address how robots can convey their FoV to align with people’s mental models, we first use a similar taxonomy (Figure 2) in our prior work (under review) that classifies FoV indicators along a spectrum from **egocentric** (robot-oriented) to **allocentric** (task-oriented) in four connected areas: Eye, head, transition, and task space. *Egocentric* designs focus on the robot, which possesses the property of FoV, directly influencing the robot’s ability to perceive its surroundings. Expanding rightward, transition space designs extend from the robot into its operating environment, and gradually transit to the task space, bridging the gap between the robot and the environment. Finally, in *allocentric* space, designs are task-oriented. They are independent of the robot’s physical body and are placed directly in the robot’s operational environment to reflect the robot’s FoV. Below we describe three baseline conditions and five designs grounded in this taxonomy.

Three baselines are **(1) Baseline**: Original robot with no indicator; **(2) Screen**: Robot showing its camera feeds; **(3) Voice**: Robot announces its FoV: “My field of view is 54.4°.”

Five designs are **(4) Glasses**: A static visual cue focusing on the robot’s eye space, which draws inspiration from people wearing glasses in life. The glasses’ legs are widened to block areas outside the robot’s FoV. **(5) Near-Eye Hands**: A static gesture-based approach focusing on the robot’s head space, the robot raises its hands to the sides of its eyes to visually indicate the range of its FoV. **(6) Motion**: A dynamic gesture focusing on bridging the gap between head space and task space. Robot starts moving its hands from the near-eye hands position, then extends its arms to the task space, revealing the vision range. **(7) Extended Arms**: A static gesture-based approach, the robot extends both arms forward to visually define its operational workspace, linking the robot

with its direct task environment, revealing the vision range the robot can see. **(8) Spotlight Scene**: A static task-space design inspired by spotlights in life, projects lights direct user attention to what lies within the robot’s vision range. A projector will be mounted on the robot’s head to dynamically project the spotlight effect onto the task space.

### IV. HYPOTHESIS

To explain a robot’s vision capability, we seek to answer a key research question: **What design strategies can most effectively help users understand a robot’s FoV while maintaining trust and positive perceptions of the robot?** Particularly, we aim to test two groups of hypotheses.

**1. Objective Outcomes**: We believe designs closer to the allocentric space will bring task-related objective benefits, as they provide direct cues to reveal the FoV in the task space. Specifically, we hypothesize that they will

**H1 – Improved Effectiveness in Understanding**: Improve people’s understanding of the robot’s FoV, to be measured by error rates and a 7-item Likert scale question on how understandable the indicator is.

**H2 – Increased Efficiency**: Improve task efficiency, to be measured by participants’ reaction times during task.

**H3 – Reduced Effort**: Lower participants’ cognitive effort, to be measured by the NASA Task Load Index.

**2. Trust and Perceptions**: We believe our proposed designs will achieve higher positive perceptions as they are more familiar designs and activate human-human interaction patterns. We also believe egocentric designs will have higher positive perceptions and trust than allocentric designs, as they leverage human-like features to create relatable experiences.

**H4 – Positive Perception**: **(a)**. Transition-space design will receive the highest positive perception scores, as it dynamically connect robot to its direct task environment, followed by head-space and eye-space static designs, with the task-space design ranked lowest. **(b)**. The proposed designs will yield higher positive perception scores than the baseline conditions. Perceptions will be measured by anthropomorphism, likability, and competence, as well as a preference question.

**H5 – More Trust**: Egocentric designs will foster greater trust, measured by the Multi-Dimensional Measure of Trust.

### V. METHOD

To test the hypotheses, we design and will conduct a 1×8 within-subject study and will control ordering effects using a balanced Latin square design [34].

#### A. Apparatus

**Robot Platform**: We will use a Pepper robot [35], a 1.2m (3.9ft) tall, two-armed humanoid robot developed by Aldebaran. The robot has a common limited horizontal FoV of 54.4° [1], and its hands can grasp light objects.

**Task Environment**: We will use the FetchIt kitting task environment [36], originally designed as a challenge task for the FetchIt Mobile Manipulation competition at ICRA 2019. In this scenario, a robot navigates to different stations to collect

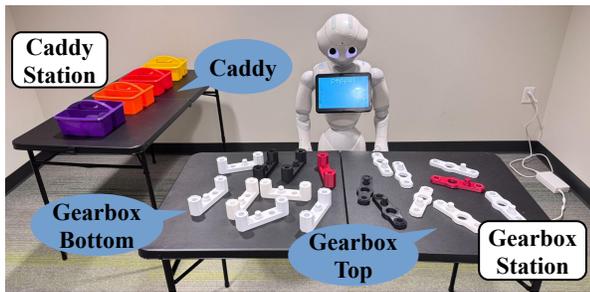


Fig. 3: Task environment adapted from the FetchIt competition [36]: Robots place parts in caddies for human assembly.

gearbox parts, place them in a caddy, and deliver the caddy to the assembly table for assembly workers to assemble a gearbox. Replenishment workers will replenish parts.

### B. Motivating Replenishment Scenario, Task, and Procedure

We investigate a scenario where a replenishment worker restocks gearbox bottoms at the gearbox station while an assembly worker assembles gearboxes at the caddy station (Figure 3). In this scenario, the replenishment worker places the first gearbox top out of the robot’s FoV. The assembly worker sees the same part missing from the caddy—a common failure scenario reported in the FetchIt task due to irregular caddy with sections easily occluded [37, 38] and asks: “Hey Pepper, can you pass me the gearbox top?” Pepper proactively responds by explaining its incapability via assumption checkers [39, 40, 38]: “I do not perceive any gearbox top so I will not be able to grasp a gearbox top,” and attempts to recover from its failure by scanning the station by moving its head. Upon spotting the gearbox top, it says: “I perceive a gearbox top so I will grasp the gearbox top.”

Upon detecting the replenishment worker’s confusion that the robot could not see the gearbox bottom that the worker just placed, Pepper explains: “Due to my limited vision capability, I can only perceive objects within my field of view. However, I can provide several ways to help you understand my horizontal field of view. Before we proceed, please place all the other objects you need to deliver on the table.”

In the experiment, participants will first read a consent form. Upon agreement, they will complete a demographic questionnaire. Then, participants will act as replenishment workers, replenishing the parts at the gearbox station, and the experimenter will act as the assembly worker. After Pepper speaks, participants will place ten gearbox tops and ten gearbox bottoms on the gearbox station. Then Pepper will try to indicate its field of view through all of the eight conditions in the assigned Latin square order in this within-subjects experiment. Specifically, the robot’s base will rotate randomly to prevent participants from using prior knowledge of the FoV thus influencing their evaluation. After each condition, they will guess which parts they believe Pepper can see as quickly as possible and finish the questionnaires.

For implementation, we will 3D model glasses and the mount of the projector in spotlight scene in SolidWorks and

then 3D print them. For gestures, we will make sure the arms accurately reflect its FoV by checking Pepper’s camera feed. We will program Pepper with Choregraphe [41], an application in part to design robot behaviors and speech.

### C. Data Collection and Measures

To measure the constructs in our hypotheses, we will use two objective metrics and five subjective metrics.

For objective measures, **error rate** will be calculated by wrong guesses/all guesses. There are two cases where participants can make a wrong guess: (1) The part is within the robot’s FoV but they guess it is outside; (2) The part is outside the robot’s FoV but they guess it is within. **Reaction time** will be calculated as the duration from when participants start observing the design to when they start to fill out the questionnaire with what parts they believe are within FoV.

For subjective measures, we will use the NASA Task Load Index [42, 43] to measure **cognitive effort**, with both the load survey and its weighting component to calculate a weighted score. **Anthropomorphism** and **likability** will be measured using the Godspeed Anthropomorphism and Likability scales [44]. **Competence** will be measured using the ROSAS scale [45]. **Trust** will be measured using the widely cited Multi-Dimensional Measure of Trust (MDMT) [46, 47], which captures both the performance and moral aspects of trust.

### D. Data Analysis

We will use a Bayesian analysis framework [48] to measure evidence for or against both the null ( $\mathcal{H}_0$ ) and alternative ( $\mathcal{H}_1$ ) hypotheses. This method uses the Bayes Factor (BF) to compare how likely the observed data are under  $\mathcal{H}_1$  (presence of an effect) versus  $\mathcal{H}_0$  (absence of an effect). For instance,  $BF_{10}=8$  indicates that the data are eight times more likely to occur under  $\mathcal{H}_1$  than  $\mathcal{H}_0$ , thereby supporting  $\mathcal{H}_1$ . We will also use a credible interval (CI) instead of Frequentist’s confidence interval. A credible interval directly states that there is a  $\alpha\%$  probability that the parameter falls within the interval.

To interpret our Bayes Factor results, we will use the widely adopted discrete classification scheme by Lee and Wagenmakers [49]. For evidence favoring  $\mathcal{H}_1$ ,  $BF_{10}$  is deemed anecdotal (inconclusive) when  $BF_{10} \in (1,3]$ , moderate when  $BF_{10} \in (3,10]$ , strong when  $BF_{10} \in (10,30]$ , very strong when  $BF_{10} \in (30,100]$ , and extreme when  $BF_{10} \in (100, \infty)$ . Anecdotal evidence is considered inconclusive while others are conclusive.

## VI. CONCLUSION

In this workshop paper, we aim to leverage familiar experiences to bridge the gap between user mental models and a robot’s vision capabilities while maintaining task performance and subjective experience. We introduce five indicator designs to be evaluated in a motivating replenishment scenario with three baseline conditions. We will conduct a user study to evaluate task performance and participants’ subjective experiences. With future results, we hope to provide insights for robot designers to better design a robot through effective indicators or gestural “open-box tutorials” that familiarize first-time users with the robot’s vision capabilities.

## REFERENCES

- [1] Aldebaran, “Pepper- technical specifications,” 2022. [Online]. Available: <https://support.aldebaran.com/support/solutions/articles/80000958735-pepper-technical-specifications>
- [2] A. Robotics, “Nao v6 video documentation,” 2014. [Online]. Available: [http://doc.aldebaran.com/28/family/nao\\_technical/video\\_naov6.html](http://doc.aldebaran.com/28/family/nao_technical/video_naov6.html)
- [3] M. Wise, M. Ferguson, D. King, E. Diehr, and D. Dymesich, “Fetch and freight: Standard platforms for service robot applications,” in *Workshop on autonomous mobile service robots*, 2016, pp. 1–6.
- [4] F. Robotics, “Robot hardware overview — fetch & freight research edition melodic documentation,” 2024. [Online]. Available: [https://docs.fetchrobotics.com/robot\\_hardware.html](https://docs.fetchrobotics.com/robot_hardware.html)
- [5] J. R. Wilson and A. Rutherford, “Mental models: Theory and application in human factors,” *Human factors*, vol. 31, no. 6, pp. 617–634, 1989.
- [6] Z. Han, E. Phillips, and H. A. Yanco, “The need for verbal robot explanations and how people would like a robot to explain itself,” *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 10, no. 4, pp. 1–42, 2021.
- [7] K. Fischer, “How people talk with robots: Designing dialog to reduce user uncertainty,” *Ai Magazine*, vol. 32, no. 4, pp. 31–38, 2011.
- [8] B. C. Kok and H. Soh, “Trust in robots: Challenges and opportunities,” *Current Robotics Reports*, vol. 1, no. 4, pp. 297–309, 2020.
- [9] K. Fischer, H. M. Weigelin, and L. Bodenhausen, “Increasing trust in human–robot medical interactions: effects of transparency and adaptability,” *Paladyn, Journal of Behavioral Robotics*, vol. 9, no. 1, pp. 95–109, 2018.
- [10] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser, and R. Parasuraman, “A meta-analysis of factors affecting trust in human-robot interaction,” *Human factors*, vol. 53, no. 5, pp. 517–527, 2011.
- [11] W. Zhang, W. Wong, and M. Findlay, “Trust and robotics: a multi-staged decision-making approach to robots in community,” *AI & SOCIETY*, vol. 39, no. 5, pp. 2463–2478, 2024.
- [12] M. M. De Graaf and S. B. Allouch, “Exploring influencing variables for the acceptance of social robots,” *Robotics and autonomous systems*, vol. 61, no. 12, pp. 1476–1486, 2013.
- [13] M. Salem, F. Eyssel, K. Rohlfing, S. Kopp, and F. Joubin, “To err is human (-like): Effects of robot gesture on perceived anthropomorphism and likability,” *International Journal of Social Robotics*, vol. 5, pp. 313–323, 2013.
- [14] —, “Effects of gesture on the perception of psychological anthropomorphism: a case study with a humanoid robot,” in *Social Robotics: Third International Conference, ICSR 2011, Amsterdam, The Netherlands, November 24–25, 2011. Proceedings 3*. Springer, 2011, pp. 31–41.
- [15] A. Tabrez, M. B. Luebbbers, and B. Hayes, “A survey of mental modeling techniques in human–robot teaming,” *Current Robotics Reports*, vol. 1, pp. 259–267, 2020.
- [16] B. R. Duffy, “Anthropomorphism and the social robot,” *Robotics and autonomous systems*, vol. 42, no. 3–4, pp. 177–190, 2003.
- [17] 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.
- [18] B. R. Duffy, “Anthropomorphism and robotics,” *The society for the study of artificial intelligence and the simulation of behaviour*, vol. 20, 2002.
- [19] E. Roesler, D. Manzey, and L. Onnasch, “Embodiment matters in social hri research: Effectiveness of anthropomorphism on subjective and objective outcomes,” *ACM Transactions on Human-Robot Interaction*, vol. 12, no. 1, pp. 1–9, 2023.
- [20] J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric, “The role of physical embodiment in human-robot interaction,” in *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2006, pp. 117–122.
- [21] Z. Han, Y. Zhu, A. Phan, F. S. Garza, A. Castro, and T. Williams, “Crossing reality: Comparing physical and virtual robot deixis,” in *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, 2023, pp. 152–161.
- [22] T. Fong, I. Nourbakhsh, and K. Dautenhahn, “A survey of socially interactive robots,” *Robotics and autonomous systems*, vol. 42, no. 3–4, pp. 143–166, 2003.
- [23] M. I. Ahmad, J. Bernotat, K. Lohan, and F. Eyssel, “Trust and cognitive load during human-robot interaction,” *arXiv preprint arXiv:1909.05160*, 2019.
- [24] J. Złotowski, D. Proudfoot, K. Yogeewaran, and C. Bartneck, “Anthropomorphism: opportunities and challenges in human–robot interaction,” *International journal of social robotics*, vol. 7, pp. 347–360, 2015.
- [25] L. Onnasch and C. L. Hildebrandt, “Impact of anthropomorphic robot design on trust and attention in industrial human-robot interaction,” *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 11, no. 1, pp. 1–24, 2021.
- [26] E. Phillips, X. Zhao, D. Ullman, and B. F. Malle, “What is human-like? decomposing robots’ human-like appearance using the anthropomorphic robot (abot) database,” in *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*, 2018, pp. 105–113.
- [27] 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.
- [28] L. D. Riek, T.-C. Rabinowitch, P. Bremner, A. G. Pipe, M. Fraser, and P. Robinson, “Cooperative gestures: Effective signaling for humanoid robots,” in *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2010, pp. 61–68.
- [29] J. Xu, J. Broekens, K. V. Hindriks, and M. A. Neerinx, “Robot mood is contagious: effects of robot body language in the imitation game,” in *AAMAS*, 2014, pp. 973–980.
- [30] A. Sauppé and B. Mutlu, “Robot deictics: How gesture and context shape referential communication,” in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 2014, pp. 342–349.
- [31] A. D. Dragan, S. Bauman, J. Forlizzi, and S. S. Srinivasa, “Effects of robot motion on human-robot collaboration,” in *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*, 2015, pp. 51–58.
- [32] A. E. Block, H. Seifi, O. Hilliges, R. Gassert, and K. J. Kuchenbecker, “In the arms of a robot: Designing autonomous hugging robots with intra-hug gestures,” *ACM Transactions on Human-Robot Interaction*, vol. 12, no. 2, pp. 1–49, 2023.
- [33] A. Huang, A. Ranucci, A. Stogsdill, G. Clark, K. Schott, M. Higger, Z. Han, and T. Williams, “(gestures vaguely): The effects of robots’ use of abstract pointing gestures in large-scale environments,” in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024, pp. 293–302.
- [34] A. D. Keedwell and J. Dénes, *Latin Squares and Their Applications: Latin Squares and Their Applications*. Elsevier, 2015.
- [35] A. K. Pandey and R. Gelin, “A mass-produced sociable humanoid robot: Pepper: The first machine of its kind,” *IEEE Robotics & Automation Magazine*, vol. 25, no. 3, pp. 40–48, 2018.
- [36] F. Robotics, “Fetch robotics open source competition.” [Online]. Available: <https://web.archive.org/web/20240828140445/https://opensource.fetchrobotics.com/competition>
- [37] Z. Han, D. Giger, J. Allspaw, M. S. Lee, H. Admoni, and H. A. Yanco, “Building the foundation of robot explanation generation using behavior trees,” *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 10, no. 3, pp. 1–31, 2021.
- [38] G. LeMasurier, A. Gautam, Z. Han, J. W. Crandall, and H. A. Yanco, “Reactive or proactive? how robots should explain failures,” in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024, pp. 413–422.
- [39] A. Gautam, T. Whiting, X. Cao, M. A. Goodrich, and J. W. Crandall, “A method for designing autonomous agents that know their limits,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2022.
- [40] X. Cao, A. Gautam, T. Whiting, S. Smith, M. A. Goodrich, and J. W. Crandall, “Robot proficiency self-assessment using assumption-alignment tracking,” *IEEE Transactions on Robotics*, vol. 39, no. 4, pp. 3279–3298, 2023.
- [41] S. Robotics, *Choregraphe 2.5.10*. [Online]. Available: <https://corporate-internal-prod.aldebaran.com/en/support/pepper-naoqi-2-9/choregraphe-setup-2510-windows>
- [42] S. G. Hart, “Nasa-task load index (nasa-tlx); 20 years later,” in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, no. 9. Sage publications Sage CA: Los Angeles, CA, 2006, pp. 904–908.
- [43] NASA, “NASA TLX paper & pencil version,” <https://humansystems.arc.nasa.gov/groups/tlx/tlxpaperpencil.php>, 2019, accessed: 2024-01-22.
- [44] 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, pp. 71–81, 2009.

- [45] 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.
- [46] B. F. Malle and D. Ullman, “A multidimensional conception and measure of human-robot trust,” in *Trust in human-robot interaction*. Elsevier, 2021, pp. 3–25.
- [47] D. Ullman and B. F. Malle, “What does it mean to trust a robot? steps toward a multidimensional measure of trust,” in *Companion of the 2018 acm/ieee international conference on human-robot interaction*, 2018, pp. 263–264.
- [48] 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.
- [49] M. D. Lee and E.-J. Wagenmakers, *Bayesian cognitive modeling: A practical course*. Cambridge university press, 2014.