Human Preferences for Robot Eye Gaze in Human-to-Robot Handovers

Tair Faibish[1] · **Alap Kshirsagar[1,2]** · **Guy Hoffman[2]** · **Yael Edan[1]**

Abstract

This paper investigates human's preferences for a robot's eye gaze behavior during human-to-robot handovers. We studied gaze patterns for all three phases of the handover process: reach, transfer, and retreat, as opposed to previous work which only focused on the reaching phase. Additionally, we investigated whether the object's size or fragility or the human's posture affect the human's preferences for the robot gaze. A public data-set of human-human handovers was analyzed to obtain the most frequent gaze behaviors that human receivers perform. These were then used to program the robot's receiver gaze behaviors. In two sets of user studies (video and in-person), a collaborative robot exhibited these gaze behaviors while receiving an object from a human. In the video studies, 72 participants watched and compared videos of handovers between a human actor and a robot demonstrating each of the three gaze behaviors. In the in-person studies, a different set of 72 participants physically performed object handovers with the robot and evaluated their perception of the handovers for the robot's different gaze behaviors. Results showed that, for both observers and participants in a handover, when the robot exhibited *Face-Hand-Face* gaze (gazing at the giver's face and then at the giver's hand during the reach phase and back at the giver's face during the retreat phase), participants considered the handover to be more likable, anthropomorphic, and communicative of timing ($p < 0.0001$). However, we did not find evidence of any effect of the object's size or fragility or the giver's posture on the gaze preference.

Keywords Human-robot handovers · Human-robot interaction · Robot eye gaze · Human-human-handovers · Non-verbal communication

Tair Faibish and Alap Kshirsagar contributed equally to this work.

This work was part of Tair Faibish Engineering Final Project and MSc thesis.

✉ Tair Faibish
tairse@post.bgu.ac.il

Alap Kshirsagar
ak2458@cornell.edu

Guy Hoffman
hoffman@cornell.edu

Yael Edan
yael@bgu.ac.il

[1] Department of Industrial Engineering and Management and the ABC Robotics Initiative, Ben-Gurion University of the Negev, Be'er Sheva, Israel

[2] Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, USA

1 Introduction

People frequently hand over objects to others or receive objects from others. Robots in domestic and industrial environments will be expected to perform such handovers with humans. For example, collaborative manufacturing (e.g., assembly), surgical assistance, household chores, shopping assistance, and elder care involve object handovers between the actors. In this work, we investigate where should a robot direct its gaze when it is receiving an object from a human.

A handover typically consists of three phases [1]: a reach phase in which both actors extend their arms towards the handover location, a transfer phase in which the object is transferred from the giver's hand to the receiver's hand, and a retreat phase in which the actors exit the interaction. These phases involve both physical and social interactions consisting of hand movements, grasp forces, body postures, verbal cues and eye gazes.

Most of the research on human-human and human-robot handovers has focused on arm movement and grasping in handovers, with only a few works that studied the social inter-

actions. Eye gaze is an important non-verbal communication mode in human-human and human-robot interactions, and it has been shown to affect the human's subjective experience of human-robot handovers [2–6]. However, except for our previous work [6], all of the prior studies of gaze behaviors in handovers considered only the robots-as-givers scenario i.e. robot-to-human handovers. Human-to-robot handovers are equally important with many applications in various domains. Some examples include a collaborative assembly task in which the robot receives parts from the human or an elder care robot that takes an empty tray from an older adult after giving him/her food.

In our previous work [6], we studied the effects of robot head gaze during the reach phase of human-to-robot handover. Results revealed that observers of a handover perceived a *Face-Hand transition* gaze, in which the robot initially looks at the giver's face and then at the giver's hand, as more anthropomorphic, likable and communicative of timing compared to continuously looking at the giver's face (*Face* gaze) or hand (*Hand* gaze). Participants in a handover perceived *Face* gaze or *Face-Hand transition* gaze as more anthropomorphic and likable compared to *Hand* gaze. However, these results were limited to a specific scenario where the giver stood in front of the robot and handed over a specific object (a plastic bottle) to the robot. Furthermore, the robot's gaze behaviors were studied only in the reach phase of the handover.

The goal of this paper is to expand and generalize the findings from our previous work. Here, we study the human's preference for robot gaze behaviors in human-to-robot handovers for all three phases of a handover for four different object types and two giver postures. Also, we use eye gaze instead of head gaze since it is more common. We also contribute to the literature on human-human handovers by identifying common gaze behaviors of humans in handovers.

2 Related Work

2.1 Human-to-Robot Handovers

Researchers have studied human-to-robot handovers to understand human preferences for robot behaviors in the approach, reach and transfer phases of handovers. In this work, we use the findings from these studies to design the robot's handover trajectory and configuration.

Investigation of the interaction of a robot handing over a can to a human [7] revealed that the preferred interpersonal distance between the human and the robot is within personal distance (0.6m - 1.25m), suggesting that people may treat robots similar to other humans. Previous research also showed that subjects understood the robot's intention during a handover by the robot's approaching motion, even without

prior knowledge in robotics or exact directions [8]. Furthermore, Cakmak et al. [9] found that handover intent also relies on handover poses, and inadequately designed handover poses might fail to convey the handover intent. Their recommendation was to create the handover pose distinct from the object holding pose. They also suggested that the best handover intent is conveyed by an almost extended arm [10]. A study of effect of participant's previous encounters with robots on human-robot handovers showed that naive users, as opposed to experienced ones, expect the robot to monitor the handover visually, rather than merely use the force sensor [11]. A study of the impact of repeated handover experiments on the robot's social perception [12] showed that participants' emotional warmth towards the robot and comfort were improved by repeated experiments.

2.2 Gaze in Handovers

There is surprisingly little work on gaze behaviors in human-to-human handovers or object passing tasks [6]. Flanagan et al. [13] investigated gaze behavior in a block stacking task. Contrary to previous assumptions, they showed that human gazes were not reactive during the task i.e. people did not focus on the gripped object or the object in movement. Instead, human gazes were found to be predictive; their gazes focused on the object's final destinations. Investigation of the discriminative features that represent the intent to start a handover revealed that mutual gaze during the task, which is often considered crucial for communication, was not a critical discriminative feature [14]. Instead, givers' initiation of a handover was better predicted using asynchronous eye gaze exchange.

In a human-to-human handover study of a water bottle [2], it was found that the givers exhibited two types of gaze behaviors: shared attention gaze and turn-taking gaze. In shared attention gaze, the giver looked at the handover location, and in turn-taking gaze, the giver initially looked at the handover location and then at the receiver's face. In our prior work [6], we found that the most common gaze behavior for both the giver and the receiver was to continuously look at the other person's hand during the reach phase of a handover. Receivers exhibited this behavior almost twice as frequently as the givers. However, our prior work studied the gaze behaviors only in the reach phase of human-to-human handovers. To the best of our knowledge, there is no prior work that studies both the giver's and the receiver's gaze in all three phases of the handover process: reach, transfer, retreat. This gap is addressed in Sect. 3.3.

Past research revealed that robot gaze affects the subjective experience and timing of robot-to-human handovers [2–5,15]. A "turn-taking gaze" in which the robot switched its gaze from the handover location to the receiver's face halfway through the handover was favoured [2]. In a follow-

up study, results revealed that the participants reached for the object sooner when the robot exhibited a "face gaze" i.e. continuously looked at receiver's face, as opposed to a shared attention gaze [3]. Fischer et al. [4] assigned a robot to retrieve parts according to participants' directions and compared two robot gaze behaviors during this task. They found that when the robot looked at the person's face instead of looking at it's own arm, participants were quicker to engage with the robot, smiled more often, and felt more responsible for the task. In a similar study, [5] it was found that when the robot looked at the participant's face while approaching them with an object, it significantly increased the robot's social presence, perceived intelligence, animacy, and anthropomorphism. Admoni et al. [15] used the robot's gaze behavior to instruct the human to place the handed-over object at a specific location. They showed that delays in the robot's release of an object draws human attention to the robot head and gaze and increases the participants' compliance with the robot's gaze behavior. In our prior work [6], we found that observers of a human-to-robot handover preferred a transition gaze in which the robot initially looked at their face and then at their hand during the reach phase. For participants in human-to-robot handovers, a face gaze was almost equally preferred as a transition gaze, though the evidence was statistically weaker.

A common limitation of these prior studies is that they do not investigate the effect of the object or the human's posture on the human's preference of robot gaze. Therefore, in the current study, as described in Sections 4–5, human preferences towards robot gaze behaviors in human-to-robot handovers for four different object types and two human postures is compared.

3 Methodology

3.1 Overview

This research aims to investigate human preferences for robot gaze behaviors in human-to-robot handovers for all three phases of the handover process (reach, transfer and retreat). To obtain possible options for robot gaze behaviors we first studied gaze behaviors in human-to-human handovers. A data set of videos of human-human handovers was analyzed, and the most common gaze behaviors of receivers were identified. Informed by this analysis, we conducted two user studies of the robot's gaze while receiving the object from the human in different situations. We investigated whether different object types or giver's postures affect the human preferences of robot gaze in human-to-robot handovers.

3.2 Hypotheses

The research hypotheses are:

- *H1*: People prefer certain robot gaze behaviors over others in terms of likability, anthropomorphism and timing communication.
- *H2*: Object size affects the user's ratings of the robot's gaze in a human-to-robot handover.
- *H3*: Object fragility affects the user's ratings of the robot's gaze in a human-to-robot handover.
- *H4*: User's posture (standing and sitting) affects the user's ratings of the robot's gaze in a human-to-robot handover.
- *H5*: Observers of a handover and participants in a handover have different preference ratings of the robot's gaze in a human-to-robot handover.

H1 is motivated by prior work which found evidence for different user preference ratings for robot gaze behaviors. We do not have a-priori hypothesis about the preference order of gaze behaviors. **H2** and **H3** are based on the intuition that the object's size and fragility could affect the preferred gaze behavior of a receiver. For example, when receiving large or fragile objects, the robot could be expected to convey attentiveness by looking at the giver's hand, whereas, when receiving small or non-fragile objects, the robot could be better off looking at the giver's face to convey friendliness. **H4** is based on the intuition that a standing giver may have different preferred gaze behavior of a receiver than a sitting giver. For example, a standing person could like the robot gaze at their face as their eyes are at the same level, whereas a sitting person could feel uncomfortable with the robot gazing down at their face. **H5** results from our previous finding that observers of a handover and participants in a handover had different preference ratings of robot gaze behaviors in the reach phase [6]. This research examines whether this holds true for robot gaze behaviors in all three phases of a handover and for handovers with different object types and giver postures.

3.3 Analysis of Gaze in Human-Human Handovers

We analyzed gaze behaviours in human-to-human handovers by annotating all three phases of each handover in a public dataset of human-human handovers [16], similar to our previous work [6]. A frame-by-frame video encoding was performed followed by annotating the giver's and receiver's gaze locations in each phase in each frame with the following discrete variables {G: Giver, R: Receiver}[1]:

1) **G's gaze:** R's face/R's hand/Own Hand/Other
2) **G's phase:** Reach/Transfer/Retreat
3) **R's gaze:** G's face/G's hand/Own Hand/Other
4) **R's phase:** Reach/Transfer/Retreat

---

[1] The annotations are available at: https://github.com/alapkshirsagar/handover-gaze-annotations/.
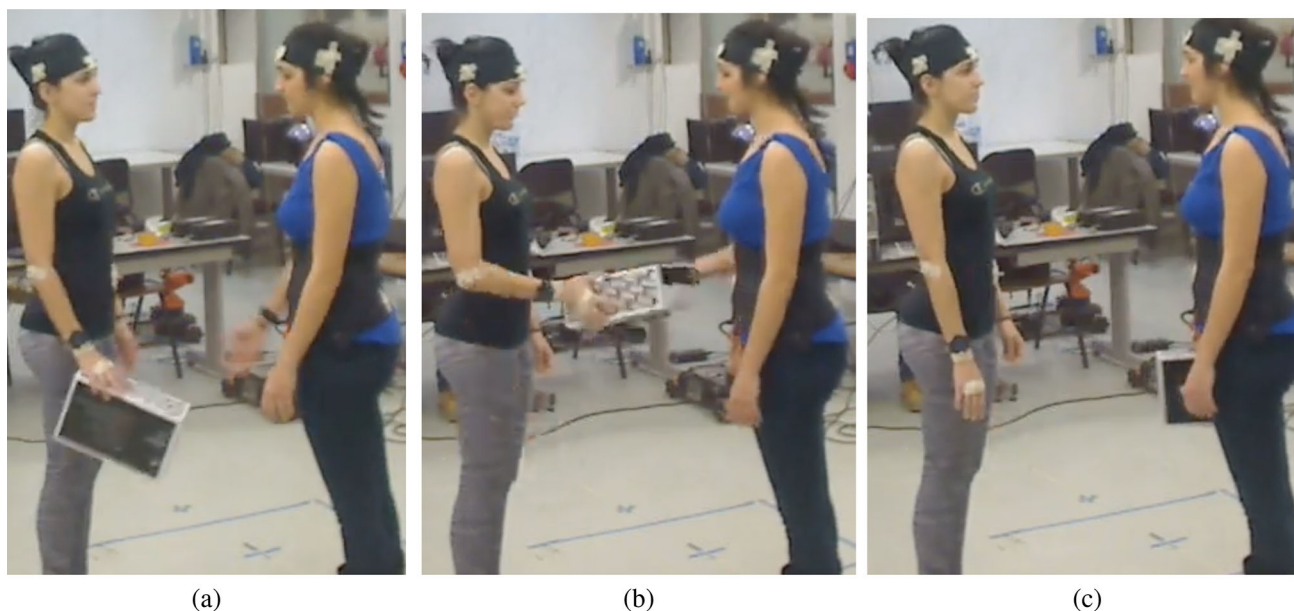
Fig. 1 Examples of gaze annotations of the human-human handovers dataset [16]. On the left is the giver and on the right the receiver: **a** Reach phase : The giver is gazing at the other's face while the receiver is gazing at the other's hand, **b** Transfer phase : Both the giver and receiver are gazing at the other's hand, **c** Retreat phase: Both the giver and the receiver are gazing at the other's face
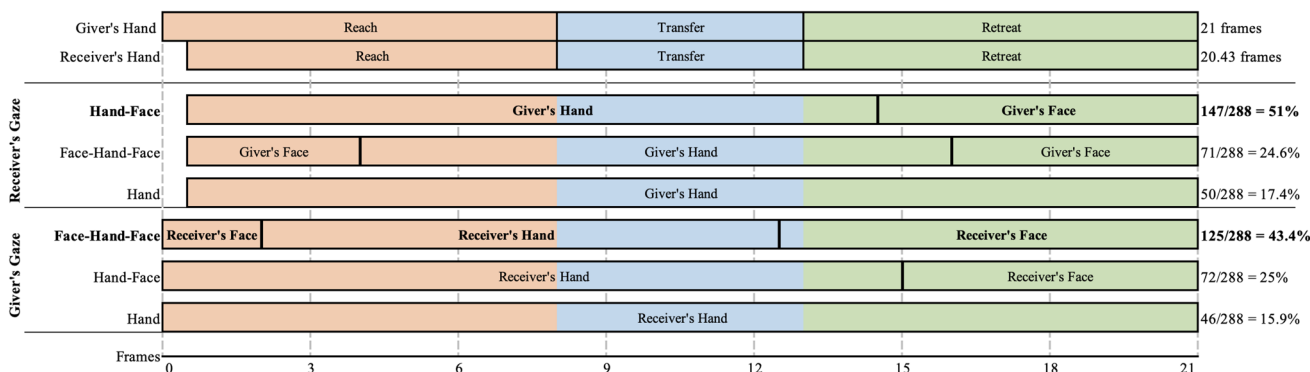

Fig. 2 Analysis of gaze behaviors in the reach, transfer and retreat phases of human-human handovers. Time flows left to right. Background colors (labeled on top two rows) correspond to each phase of a handover: red: reach; blue: transfer; green: retreat. The bottom six rows show one handover behavior each, three for the receiver and three for the giver. Boundaries correspond to average length of each phase. Prevalence of each behavior is noted at the right edge of the row. Givers and receivers have dissimilar frequently observed gaze behaviors

Figure 1 shows some examples of gaze annotations in the three phases of handovers. The analysis (Fig. 2) revealed that the most common gaze behaviors employed by people during handovers are:

1) *Hand-Face* gaze: The person continuously looks at the other person's hand during the reach and the transfer phases, and then looks at the other person's face during the retreat phase. The transition from hand to face happens slightly after the beginning of the retreat phase. More than 50% of receivers showed this behavior, whereas, only 25% of the givers in those videos exhibited this behavior.

2) *Face-Hand-Face* gaze: During the reach phase, the person initially looks at the other person's face and then at the other person's hand. They then continue looking at the other person's hand during the transfer phase. Finally they look at the other person's face during the retreat phase. The transition from face to hand occurs halfway through the reach phase, while the transition from hand to face occurs halfway through the retreat phase. More than 40% of givers exhibited this gaze, whereas only 25% of receivers did.

3) *Hand* gaze: Continuously looks at the other person's hand. The least frequent gaze, only 17.4% of receivers and 15.9% of givers showed this behavior.

3.4 Human-Robot Handover Studies

Two within-subject studies were conducted, a video study and an in-person study. The video study aimed to investigate an observer's preferences of robot gaze behaviors, whereas the in-person study aimed to investigate a giver's preferences of robot gaze behaviors.

A total of 144 undergraduate industrial engineering students participated in the experiment (72 in each study) and were compensated with one bonus point to their grade in a course for their participation. The average participation time was about 25 minutes. In the video study, there were 34 females and 38 males aged 23-29. In the in-person study, there were 36 females and 36 males aged 23-30. The study design was approved by the Human Subjects Research Committee at the Department of Industrial Engineering and Management, Ben-Gurion University of the Negev.

The following three gaze behaviors were implemented on a Sawyer cobot based on insights from the human-human handover analyses:

i. *Hand-Face* gaze: The robot's eyes continuously looked in the direction of the giver's hand during the reach and transfer phases. After the robot started to retreat, the eyes transitioned to look at the giver's face. Both the hand gaze and the face gaze were programmed manually to fixed locations.

ii. *Face-Hand-Face* gaze: The robot's eyes looked at the giver's face during the reach phase, giver's hand during the transfer phase and giver's face during the retreat phase.

iii. *Hand* gaze: The robot's eyes continuously looked in the direction of the giver's hand.

Given that the human gaze behavior was tied to the handover phase, as described above, we did not use fixed timings for the robot trajectory. Instead, the robot was programmed to use sensor information to initiate the handovers and gaze behaviors depending on the phase of the handover. The robot arm was programmed to reach a predefined position once the giver started the handover which was detected using a range sensor. The robot's gripper was equipped with an infrared proximity sensor, and it grasped the object when the object was close enough. The robot retreated to its home position after grasping the object. The robot was programmed in the Robot Operating System (ROS) environment with Rethink Robotics' Intera software development kit (SDK). The sensors were interfaced with the robot using an Arduino micro-controller.

Figure 3a shows a snapshot of a video recording illustrating the experimental setup.[2]
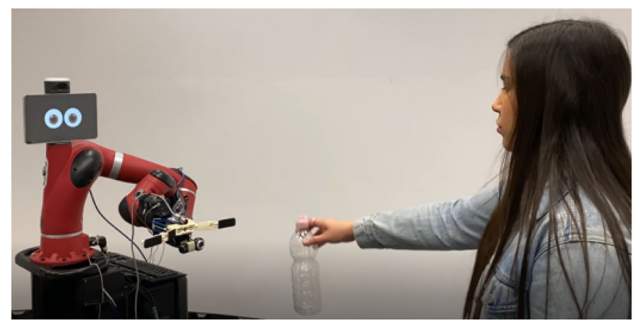
(a)
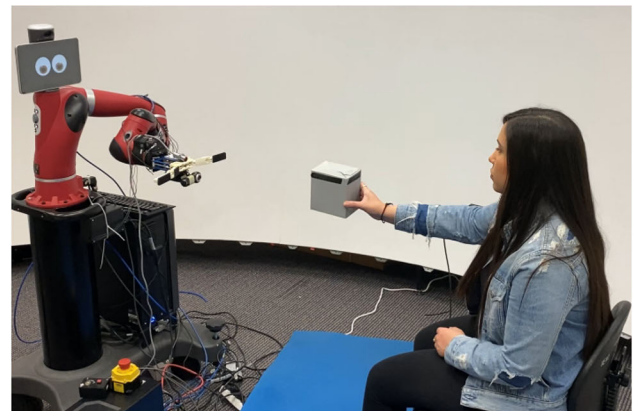
(b)
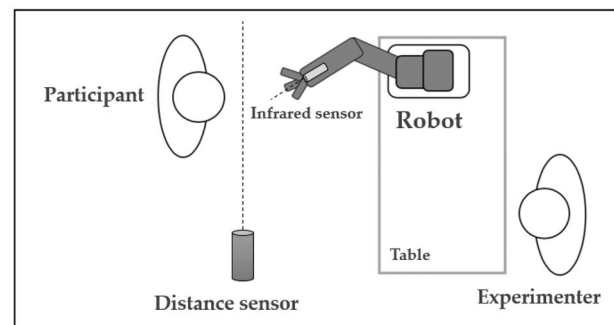
(c)

Fig. 3 Experimental Setup: Video frames of an actor handing over an object to the robot, used in the video study: **a** "Standing" posture **b** "Sitting" posture **c** Diagram of the setup for the in-person study

4 Video Study of Human-to-Robot Handovers

4.1 Experimental Procedure and Evaluation

The study was conducted remotely, and each participant received links to the videos, electronic consent form, and online questionnaires with study instructions. After signing the consent form and reading the instructions, they completed a practice session followed by 12 study sessions. Each ses-

---

sion included one of the six pairing of the gaze patterns listed in Table 2, for a single condition out of the three listed in Table 1³. So that each participant watched all six pairs of gaze patterns twice, one for condition *a* and one for condition *b*. To reduce the recency effect of participants forgetting the previous conditions counterbalanced pairwise comparisons were performed instead of three-way comparisons. All six pairwise comparisons were combined into a ranked ordered list of three gaze patterns [18]. In each session, they watched two handover videos, consecutively. The different objects and postures used in the experiment are shown in Figs. 4 and 3 respectively.

The instructions at the start of the experiment, as well as the caption for each video, stated that participants should pay close attention to the robot's eyes in the video. After every two videos, the participants were asked to fill out a questionnaire which collected subjective measures as detailed below. The questionnaire was identical to the one used in our previous study [6] and in Zheng et al.'s study [3]. Questions 1 and 2 measure the metric *likability* (Cronbach's $\alpha = 0.83$). Questions 3 and 4 measure the metric *anthropomorphism* (Cronbach's $\alpha = 0.91$). Question 5 measures the metric *timing communication*.

1) Which handover did you like better? (1st or 2nd)
2) Which handover seemed more friendly? (1st or 2nd)
3) Which handover seemed more natural? (1st or 2nd)
4) Which handover seemed more humanlike? (1st or 2nd)
5) Which handover made it easier to tell when, exactly, the robot wanted the giver to give the object? (1st or 2nd)
6) Any other comments (optional)

4.2 Experimental Design

The experiment was designed as a between-within experiment, using likability, anthropomorphism, timing commu-

---

³ To represent objects of different fragility a plastic bottle and a glass bottle were used. In order to examine people's perception about the fragility of these objects, we conducted an online survey. This survey was conducted post experiment based on reviewers' feedback. A total of 24 participants responded to the survey. The participants were undergraduate students from the Department of Industrial Engineering and Management at Ben-Gurion University, similar to the students who participated in our video and in-person experiments. The participants were told that this study deals with object handovers between a human and a robot.
The survey included 10 pictures of objects, made from different materials. The plastic bottle and the glass bottle used in our experiment were among these objects. Each picture was followed by a yes or no question: "Do you perceive this object to be fragile?". Results revealed that all of the 24 participants perceived the plastic bottle to be non-fragile. 23 out of 24 participants perceived the glass bottle to be fragile. Additionally, when asked the same question for three other different plastic and glass bottles, 24 participants denoted the plastic bottles as non-fragile and 23 denoted the glass bottles as fragile. Details about this survey are available in [17]. This supports our decision to choose plastic and glass bottles to represents objects of different fragility.
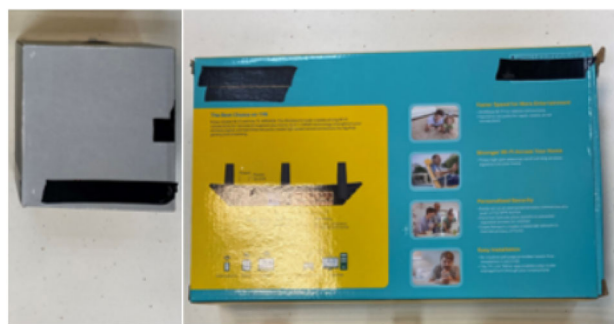
(a)

(b)

Fig. 4 The objects used in the experiments: **a** Object size (small box and large box), **b** Object fragility (plastic bottle and glass bottle)

Table 1 Study Conditions (24 participants per condition)

Condition 1: Object Size	*a*. Small Box
	b. Large Box
Condition 2: Object fragility	*a*. Plastic Bottle
	b. Glass Bottle
Condition 3: User's Posture	*a*. Sitting
	b. Standing

Table 2 Six pairings of the three gaze patterns and their reverse order for each object or posture. Each participant experienced two versions (a/b of a single condition) of these pairings, for a total of 12 pairings

First Handover	Second Handover
Hand-Face	Face-Hand-Face
Hand-Face	Hand
Face-Hand-Face	Hand
Face-Hand-Face	Hand-Face
Hand	Hand-Face
Hand	Face-Hand-Face

nication as the dependent variables. The participants were divided into three groups of 24 participants. Each group per-

Table 3 Combined preferences of gaze behaviors in the video study for the small and large object conditions

		Small Object					Large Object				
		Hand-Face	Face-Hand-Face	Hand	a_i	P_i	Hand-Face	Face-Hand-Face	Hand	a_i	P_i
Likability	Hand-Face	0	5.25	21	26.25	0.2	0	5.75	21	26.75	0.25
	Face-Hand-Face	18.75	0	23.5	42.25	0.77	18.25	0	21	39.25	0.70
	Hand	3	0.5	0	3.5	0.03	3	3	0	6	0.05
Anthropomorphism	Hand-Face	0	5	20.5	25.5	0.21	0	7.75	20.5	28.25	0.28
	Face-Hand-Face	19	0	22.5	41.5	0.75	16.25	0	20.75	37	0.65
	Hand	3.5	1.5	0	5	0.04	3.5	3.25	0	6.75	0.07
Timing Communication	Hand-Face	0	7.5	19	26.5	0.29	0	6	20	26	0.31
	Face-Hand-Face	16.5	0	21.5	38	0.63	18	0	17.5	35.5	0.57
	Hand	5	2.5	0	7.5	0.08	4	6.5	0	10.5	0.12

Table 4 Combined preferences of gaze behaviors in the video study for the non-fragile object and fragile object conditions

		Non-Fragile Object					Fragile Object				
		Hand-Face	Face-Hand-Face	Hand	a_i	P_i	Hand-Face	Face-Hand-Face	Hand	a_i	P_i
Likability	Hand-Face	0	6.25	19.5	25.75	0.27	0	7.25	21	28.25	0.32
	Face-Hand-Face	17.75	0	20.5	38.25	0.65	16.75	0	21	37.75	0.62
	Hand	4.5	3.5	0	8	0.08	3	3	0	6	0.06
Anthropomorphism	Hand-Face	0	6	18.25	24.25	0.27	0	7.5	19.5	27	0.32
	Face-Hand-Face	18	0	19	37	0.62	16.5	0	19.75	36.25	0.59
	Hand	5.75	5	0	10.75	0.11	4.5	4.25	0	8.75	0.09
Timing Communication	Hand-Face	0	7	19.5	26.5	0.34	0	7.5	19.5	27	0.35
	Face-Hand-Face	17	0	17	34	0.53	16.5	0	17	33.5	0.52
	Hand	4.5	7	0	11.5	0.13	4.5	7	0	11.5	0.13

Table 5 Combined preferences of gaze behaviors in the video study for the standing and sitting conditions

		Standing					Sitting				
		Hand-Face	Face-Hand-Face	Hand	a_i	P_i	Hand-Face	Face-Hand-Face	Hand	a_i	P_i
Likability	Hand-Face	0	4.5	21	25.5	0.23	0	8.25	19.75	28	0.35
	Face-Hand-Face	19.5	0	21	40.5	0.72	15.75	0	19.5	35.25	0.55
	Hand	3	3	0	6	0.05	4.25	4.5	0	8.75	0.10
Anthropomorphism	Hand-Face	0	7.25	20.5	27.75	0.33	0	9.25	19	28.25	0.37
	Face-Hand-Face	16.75	0	19.75	36.5	0.59	14.75	0	18.25	33	0.50
	Hand	3.5	4.25	0	7.75	0.08	5	5.75	0	10.75	0.13
Timing Communication	Hand-Face	0	6.5	20	26.5	0.31	0	7.5	18	25.5	0.30
	Face-Hand-Face	17.5	0	19	36.5	0.60	16.5	0	19.5	36	0.59
	Hand	4	5	0	9	0.09	5	4.5	0	9	0.11

formed one of the three study conditions listed in Table 1. The order of the 12 sessions were randomized and counterbalanced among the subjects.

4.3 Analysis

The participants' ratings for the likability and anthropomorphism of the gaze behaviors were measured by averaging their responses to Questions 1-2 and 3-4 respectively. The one-sample Wilcoxon signed-rank test was used to check if participants exhibited any bias towards selecting the first or the second handover. Similar to our previous work [6] and Zheng et. al's work [3], the Bradley-Terry model [19] was used to evaluate participants' rankings of the likeability, anthropomorphism and timing communication of gaze behaviors. To evaluate the hypothesis **H1**, i.e. $P_i \neq P_j \forall i \neq j$, where P_i is the probability that one gaze condition is preferred over others, the χ^2 values for each metric were computed, as proposed by Yamaoka et. al [20]:

$$B = n \sum_{i<j} log(P_i + P_j) - \sum_i a_i log P_i, \tag{1}$$

$$\chi^2 = ng(g-1)ln2 - 2Bln10, \tag{2}$$

where, $g = 3$ is the number of gaze behaviors, n is the number of participants, a_i is the sum of ratings in each row of Tables 3-7 (Appendix).

In order to examine **H2-H4**, we conducted two series of tests for each measured metric (likability, anthropomorphism and timing communication), and for each study scenario:

- Binary proportion difference tests for matched pairs [21], in which the difference between the proportion of participants who chose one gaze condition p_b over other p_c was evaluated in each study scenario. The distribution of differences $p_b - p_c$ is:

$$p_b - p_c \sim \mathcal{N}(0, \sqrt{\frac{p_b + p_c - (p_b - p_c)^2}{n}}) \tag{3}$$

where $n = 24$ is the number of participants in each scenario. The Z-score is calculated according to the following formula:

$$Z = \frac{(p_b - p_c)}{\sqrt{var(p_b - p_c)}} \tag{4}$$

A low Z-score means that the distribution of differences has zero mean with high probability.

- Equivalence tests based on McNemar's test for matched proportions [22,23], in which the proportion of participants who changed their gaze preferences in each study
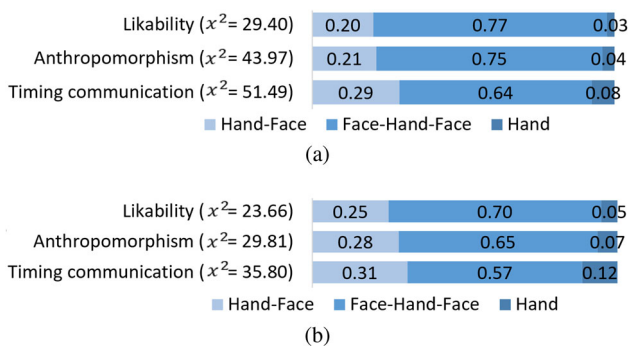
Likability (x^2= 29.40): 0.20 | 0.77 | 0.03
Anthropomorphism (x^2= 43.97): 0.21 | 0.75 | 0.04
Timing communication (x^2= 51.49): 0.29 | 0.64 | 0.08
■ Hand-Face ■ Face-Hand-Face ■ Hand
(a)

Likability (x^2= 23.66): 0.25 | 0.70 | 0.05
Anthropomorphism (x^2= 29.81): 0.28 | 0.65 | 0.07
Timing communication (x^2= 35.80): 0.31 | 0.57 | 0.12
■ Hand-Face ■ Face-Hand-Face ■ Hand
(b)

Fig. 5 χ^2 values and win-probabilities of gaze conditions in the video study for the three dependent measures: **a** Small object , **b** Large object

Likability (x^2= 28.52): 0.27 | 0.65 | 0.08
Anthropomorphism (x^2= 20.63): 0.27 | 0.62 | 0.11
Timing communication (x^2= 15.42): 0.34 | 0.53 | 0.13
■ Hand-Face ■ Face-Hand-Face ■ Hand
(a)

Likability (x^2= 33.51): 0.32 | 0.62 | 0.06
Anthropomorphism (x^2= 23.74): 0.32 | 0.59 | 0.09
Timing communication (x^2= 14.98): 0.35 | 0.52 | 0.13
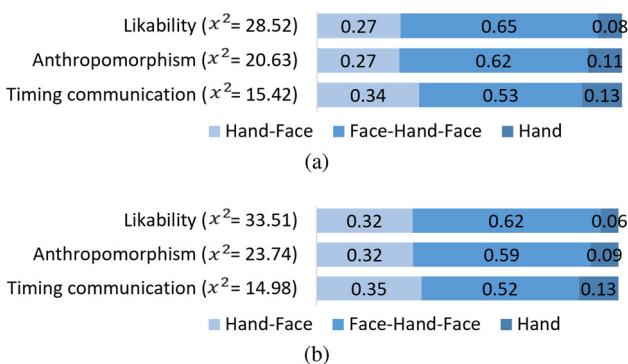■ Hand-Face ■ Face-Hand-Face ■ Hand
(b)

Fig. 6 χ^2 values and win-probabilities of gaze conditions in the video study for the three dependent measures: **a** Non-Fragile object , **b** Fragile object
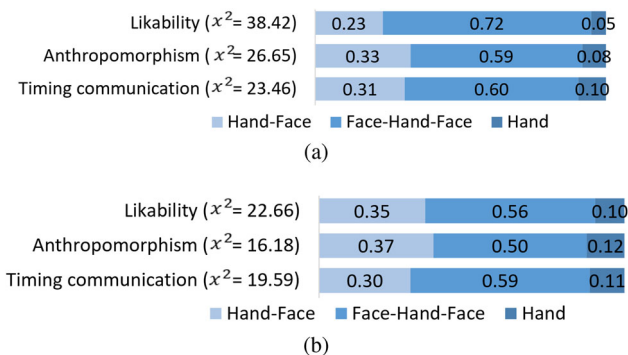
Likability (x^2= 38.42): 0.23 | 0.72 | 0.05
Anthropomorphism (x^2= 26.65): 0.33 | 0.59 | 0.08
Timing communication (x^2= 23.46): 0.31 | 0.60 | 0.10
■ Hand-Face ■ Face-Hand-Face ■ Hand
(a)

Likability (x^2= 22.66): 0.35 | 0.56 | 0.10
Anthropomorphism (x^2= 16.18): 0.37 | 0.50 | 0.12
Timing communication (x^2= 19.59): 0.30 | 0.59 | 0.11
■ Hand-Face ■ Face-Hand-Face ■ Hand
(b)

Fig. 7 χ^2 values and win-probabilities of gaze conditions in the video study for the three dependent measures: **a** Standing, **b** Sitting

scenario was compared within equivalence bounds of $\triangle = \pm 0.1$.

Table 6 Combined preferences of gaze behaviors in the in-person study for the small and large object conditions

		Small Object					Large Object				
		Hand-Face	Face-Hand-Face	Hand	a_i	P_i	Hand-Face	Face-Hand-Face	Hand	a_i	P_i
Likability	Hand-Face	0	6.5	20.75	27.25	0.29	0	7.25	20	27.25	0.31
	Face-Hand-Face	17.5	0	20.75	38.25	0.64	16.75	0	20.75	37.5	0.62
	Hand	3.25	3.25	0	6.5	0.07	4	3.25	0	7.25	0.07
Anthropomorphism	Hand-Face	0	6.5	21	27.5	0.3	0	7.5	19	26.5	0.29
	Face-Hand-Face	17.5	0	20.75	38.25	0.64	16.5	0	21.25	37.75	0.63
	Hand	3	3.25	0	6.25	0.06	5	2.75	0	7.75	0.08
Timing Communication	Hand-Face	0	7.5	22	29.5	0.37	0	8.5	18.5	27	0.34
	Face-Hand-Face	16.5	0	19	35.5	0.55	15.5	0	19	34.5	0.54
	Hand	2	5	0	7	0.08	5.5	5	0	10.5	0.12

Table 7 Combined preferences of gaze behaviors in the in-person study for the non-fragile object and fragile object conditions

		Non-Fragile Object					Fragile Object				
		Hand-Face	Face-Hand-Face	Hand	a_i	P_i	Hand-Face	Face-Hand-Face	Hand	a_i	P_i
Likability	Hand-Face	0	5.75	22	27.75	0.24	0	8	22.5	30.5	0.36
	Face-Hand-Face	18.25	0	23	41.25	0.73	16	0	21.25	37.25	0.59
	Hand	2	1	0	3	0.03	1.5	2.75	0	4.25	0.05
Anthropomorphism	Hand-Face	0	6.25	22.25	28.5	0.27	0	8	21.75	29.75	0.34
	Face-Hand-Face	17.75	0	22.75	40.5	0.70	16	0	21.5	37.5	0.61
	Hand	1.75	1.25	0	3	0.03	2.25	2.5	0	4.75	0.05
Timing Communication	Hand-Face	0	8	21.5	29.5	0.36	0	9	21	30	0.40
	Face-Hand-Face	16	0	20	36	0.57	15	0	18.5	33.5	0.50
	Hand	2.5	4	0	6.5	0.07	3	5.5	0	8.5	0.10

4.4 Results

4.4.1 Quantitative Results

To test for order effects, we checked, but did not find any bias towards selecting the first or the second handover [like: z =-0.68, $p = 0.50$; friendly: z = 1.22, $p = 0.22$; natural: z =0.20, p = 0.84; humanlike: z = 1.36, $p = 0.17$; timing communication: z =1.23, $p = 0.22$].

Tables 3 - 5 (Appendix) and Fig. 5-7 show the robot gaze preferences of the participants in terms of likability, anthropomorphism and timing communication.

Gaze conditions differ significantly in ratings (all χ^2 values are large ($p < 0.0001$)), supporting **H1**. Participants prefer the *Face-Hand-Face* transition gazes over *Hand-Face* and *Hand* gazes. *Hand* gaze is the least preferred condition.

Based on the binary proportion difference test, we did not find evidence that the proportion of observers of a handover preferring one gaze condition over the other is affected by object size (Table 9, Appendix), object fragility (Table 10, Appendix) and user's posture (Table 11, Appendix). Hypotheses **H2**, **H3** and **H4** are not supported (all p values are over 0.2).

However, based on the equivalence tests, we did not find evidence that the proportion of observers of a handover preferring one gaze condition over the other is equivalent for the two object sizes (Table 9, Appendix), object fragilities (Table 10, Appendix), or user's postures (Table 11, Appendix). Thus, hypotheses **H2**, **H3** and **H4** can also not be rejected (all p values are over 0.15).

4.4.2 Open-ended Responses

All open-ended responses are presented in [17] with major insights detailed below.

10 out of 72 participants gave at least one additional comment. Four out of the eight participants, who made *Hand-Face* gaze vs. *Face-Hand-Face* gaze comparisons, preferred *Face-Hand-Face* gaze over *Hand-Face* gaze due to the extended eye contact by the robot.

> P059 - *"As much eye contact as possible."*
> P048 - *"I preferred handover 2 (Face-Hand-Face gaze) because the robot looked more at the human"*

Two participants mentioned that they could not distinguish between *Face-Hand-Face* gaze and *Hand-Face* gaze, while two participants commented about the advantages and disadvantages of the two gaze patterns.

> P041 - *"In handover 1 (Hand-Face gaze) you could tell that the robot was ready to receive the object. However, handover 2 (Face-Hand-Face gaze) felt more human-*

ized because the robot looked at the giver's eyes right until the transfer was made".

Four out of six participants, who commented on the comparison between *Hand-Face* gaze and *Hand* gaze, preferred *Hand-Face* gaze because of the eye movement.

P008 - *"In my opinion, the change in eye movement creates a better human-robot interaction."*
P009 - *"In the second handover (Hand-Face gaze) the eye movement, gave a good indication for the communication."*

Two participants mentioned that they could not distinguish between *Hand-Face* gaze and *Hand* gaze.

Six participants commented on *Face-Hand-Face* gaze vs. *Hand* gaze comparison. All of them said that they preferred *Face-Hand-Face* gaze over *Hand* gaze.

P009 - *"At handover 2 (Face-Hand-Face gaze), the robot looked at the object precisely when it wanted to take it, so it was perceived more understandable."*
P037 - *"In my opinion video 2 (Face-Hand-Face gaze) best simulated human-like behavior out of all the videos I have seen so far."*

5 In-person Study of Human-to-Robot Handovers

In the in-person study, another set of 72 participants were asked to perform object handovers with the Sawyer robot arm in a similar setup (Fig. 3c). The robot arm and the robot eyes were programmed in the same way as the video study described in Sect. 4.

5.1 Experimental Procedure, Design and Evaluation

The experiment was conducted during the COVID-19 pandemic. Therefore, several precautions were taken. The participants were asked to wash their hands with soap when they entered and exited the lab. The equipment was sterilized before and after each participant, and the experiment room's door remained open at all times. Only one participant was allowed at a time inside the room. Both the participant and conductor of the experiment wore masks and kept at least 2 meters distance between them.

After entering the experiment room, participants signed the electronic consent form, and answered a question on a computer: *How familiar are you with a collaborative robot (such as the one shown)?* Participants ranked this question on a scale from 1 - "Not at all familiar" to 5 - "Extremely familiar". The mean familiarity with this type of robot was found to be low (M=1.49, SD = 0.60, on a scale of 1-5).

Table 8 Combined preferences of gaze behaviors in the in-person study for the standing and sitting conditions

		Standing					Sitting				
		Hand-Face	Face-Hand-Face	Hand	a_i	P_i	Hand-Face	Face-Hand-Face	Hand	a_i	P_i
Likability	Hand-Face	0	7.5	17.5	25	0.31	0	6.75	18.75	25.5	0.28
	Face-Hand-Face	16.5	0	17.75	34.25	0.54	17.25	0	20	37.25	0.62
	Hand	6.5	6.25	0	12.75	0.15	5.25	4	0	9.25	0.10
Anthropomorphism	Hand-Face	0	8.25	16.5	24.75	0.31	0	7.5	17.75	25.25	0.28
	Face-Hand-Face	15.75	0	18	33.75	0.53	16.5	0	20.5	37	0.62
	Hand	7.5	6	0	13.5	0.16	6.25	3.5	0	9.75	0.10
Timing Communication	Hand-Face	0	9	18	27	0.35	0	6.5	19	25.5	0.31
	Face-Hand-Face	15	0	18.5	33.5	0.52	17.5	0	18	35.5	0.57
	Hand	6	5.5	0	11.5	0.13	5	6	0	11	0.12

Table 9 Results of binary proportion difference test and equivalence test for matched pairs comparing small object and large object user's preferences of robot gaze in handovers. Gaze condition in bold is the preferred choice in each pairwise comparison

	Metrics	Gaze Conditions	Binary Proportion Difference Test		Equivalence Test	
			Z score	P-value	Z score	P-value
Video Study	Likability	**Hand-Face** vs. Face-Hand-Face	−0.15	0.44	0.39	0.35
		Hand-Face vs. Hand	0.00	0.50	−0.92	0.18
		Face-Hand-Face vs. Hand	0.38	0.36	0.19	0.58
	Anthropomorphism	**Hand-Face** vs. Face-Hand-Face	−0.78	0.22	0.31	0.38
		Hand-Face vs. Hand	0.00	0.50	0.80	0.21
		Face-Hand-Face vs. Hand	0.27	0.40	−0.32	0.38
	Timing communication	**Hand-Face** vs. Face-Hand-Face	0.41	0.34	−0.75	0.23
		Hand-Face vs. Hand	−0.16	0.44	−0.25	0.60
		Face-Hand-Face vs. Hand	0.65	0.26	0.04	0.52
In-Person Study	Likability	**Hand-Face** vs. Face-Hand-Face	−0.20	0.42	0.51	0.30
		Hand-Face vs. Hand	0.12	0.45	0.45	0.33
		Face-Hand-Face vs. Hand	0.00	0.50	0.96	0.17
	Anthropomorphism	**Hand-Face** vs. Face-Hand-Face	−0.27	0.40	0.59	0.28
		Hand-Face vs. Hand	0.32	0.38	0.06	0.47
		Face-Hand-Face vs. Hand	−0.08	0.47	0.27	0.39
	Timing communication	**Hand-Face** vs. Face-Hand-Face	−0.25	0.40	0.27	0.39
		Hand-Face vs. Hand	0.55	0.29	0.36	0.36
		Face-Hand-Face vs. Hand	0.00	0.50	0.30	0.38

Table 10 Results of binary proportion difference test and equivalence test for matched pairs comparing fragile object and non-fragile object user's preferences of robot gaze in handovers. Gaze condition in bold is the preferred choice in each pairwise comparison

	Metrics	Gaze Conditions	Binary Proportion Difference Test		Equivalence Test	
			Z score	P-value	Z score	P-value
Video Study	Likability	**Hand-Face** vs. Face-Hand-Face	−0.27	0.39	0.29	0.38
		Hand-Face vs. Hand	−0.24	0.41	−0.57	0.28
		Face-Hand-Face vs. Hand	−0.08	0.47	−0.92	0.18
	Anthropomorphism	**Hand-Face** vs. Face-Hand-Face	−0.41	0.34	−0.67	0.25
		Hand-Face vs. Hand	−0.20	0.42	−0.34	0.37
		Face-Hand-Face vs. Hand	−0.12	0.45	0.83	0.20
	Timing communication	**Hand-Face** vs. Face-Hand-Face	−0.13	0.45	−0.03	0.51
		Hand-Face vs. Hand	0.00	0.50	−0.33	0.37
		Face-Hand-Face vs. Hand	0.00	0.50	0.31	0.38
In-Person Study	Likability	**Hand-Face** vs. Face-Hand-Face	−0.61	0.27	−0.20	0.58
		Hand-Face vs. Hand	−0.07	0.47	0.41	0.66
		Face-Hand-Face vs. Hand	0.26	0.40	0.32	0.62
	Anthropomorphism	**Hand-Face** vs. Face-Hand-Face	−0.47	0.32	0.43	0.33
		Hand-Face vs. Hand	0.08	0.47	0.56	0.71
		Face-Hand-Face vs. Hand	0.19	0.43	−0.23	0.41
	Timing communication	**Hand-Face** vs. Face-Hand-Face	−0.24	0.41	−0.03	0.51
		Hand-Face vs. Hand	0.08	0.47	−0.14	0.44
		Face-Hand-Face vs. Hand	0.24	0.41	−0.66	0.25

Table 11 Results of binary proportion difference test and equivalence test for matched pairs comparing sitting and standing user's preferences of robot gaze in handovers. Gaze condition in bold is the preferred choice in each pairwise comparison

	Metrics	Gaze Conditions	Binary Proportion Difference Test		Equivalence Test	
			Z score	P-value	Z score	P-value
Video Study	Likability	**Hand-Face** vs. Face-Hand-Face	−1.08	0.15	−0.63	0.26
		Hand-Face vs. Hand	0.20	0.42	−0.04	0.52
		Face-Hand-Face vs. Hand	0.24	0.41	−0.05	0.52
	Anthropomorphism	**Hand-Face** vs. Face-Hand-Face	−0.49	0.31	0.71	0.24
		Hand-Face vs. Hand	0.24	0.41	−0.65	0.74
		Face-Hand-Face vs. Hand	0.24	0.41	−0.60	0.72
	Timing communication	**Hand-Face** vs. Face-Hand-Face	−0.27	0.40	0.44	0.33
		Hand-Face vs. Hand	0.33	0.37	0.16	0.43
		Face-Hand-Face vs. Hand	−0.08	0.47	−0.21	0.58
In-Person Study	Likability	**Hand-Face** vs. Face-Hand-Face	0.20	0.42	−0.28	0.39
		Hand-Face vs. Hand	−0.21	0.42	−0.51	0.31
		Face-Hand-Face vs. Hand	−0.37	0.36	0.72	0.24
	Anthropomorphism	**Hand-Face** vs. Face-Hand-Face	0.19	0.43	−0.20	0.42
		Hand-Face vs. Hand	−0.21	0.42	−0.39	0.35
		Face-Hand-Face vs. Hand	−0.40	0.35	0.06	0.48
	Timing communication	**Hand-Face** vs. Face-Hand-Face	0.64	0.26	−0.45	0.33
		Hand-Face vs. Hand	−0.16	0.44	0.48	0.32
		Face-Hand-Face vs. Hand	0.08	0.47	−0.32	0.38

The study instructions were given orally by the experimenter. Participants then completed a practice session followed by 12 randomly assigned study sessions. In each session, the participants performed two sequential handovers with the robot. The 12 sessions consisted of the same pairings of gaze behaviors as in the video experiment, followed by the same questionnaire questions. The only difference was in Question 5, which was "Which handover made it easier to tell when, exactly, the robot wanted you to give the object? (1st or 2nd)". The experimental design was also same as the video study.

5.2 Analysis

The hypotheses **H1-H4** were evaluated using the same procedure as described in Sect. 4.3.

To evaluate hypothesis **H5**, we conducted two series of tests for each measured metric (likability, anthropomorphism and timing communication), and for each study scenario. These tests are different from the tests for "matched pairs" which we performed for testing **H2-H4**, since for testing **H5** we need to compare two different participants' groups:

- Binary proportion difference tests for unmatched pairs [24], in which the difference between the proportion of participants who chose one gaze condition over other in each study scenario for the video p_b and in-person p_c

studies was evaluated. The distribution for the differences $p_b - p_c$ is:

$$p_b - p_c \sim \mathcal{N}(0, \sqrt{p_d(1 - p_d)(\frac{1}{n_b} - \frac{1}{n_c})}) \qquad (5)$$

where $n_b = 24$ and $n_c = 24$ are the number of participants in each scenario of the video study and in-person study respectively, and p_d is the pooled proportion calculated as follows:

$$p_d = \frac{X_b + X_c}{n_b + n_c} \qquad (6)$$

where X_b and X_c are the number of participants who preferred one gaze condition over the other (shown in Tables 3 - 8, Appendix) in the video and in-person study respectively. Then, the Z-score is calculated same as equation (4).

- Equivalence tests for unmatched proportions [25], in which the proportion of participants who chose one gaze condition over other in each study scenario for the video p_b and in-person p_c studies was tested for equivalence within the bounds of $\triangle = \pm 0.1$.
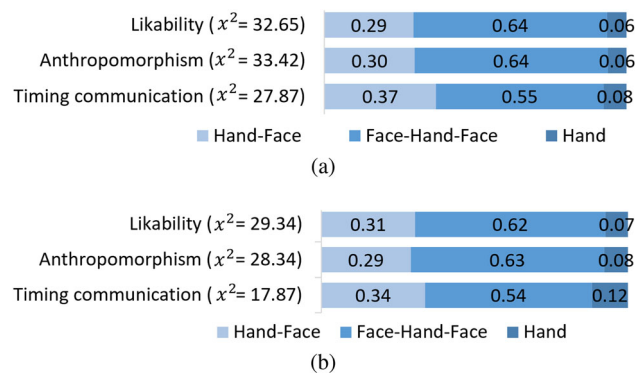
Fig. 8 χ^2 values and win-probabilities of gaze conditions in the in-person study for the three dependent measures: **a** Small object, **b** Large object
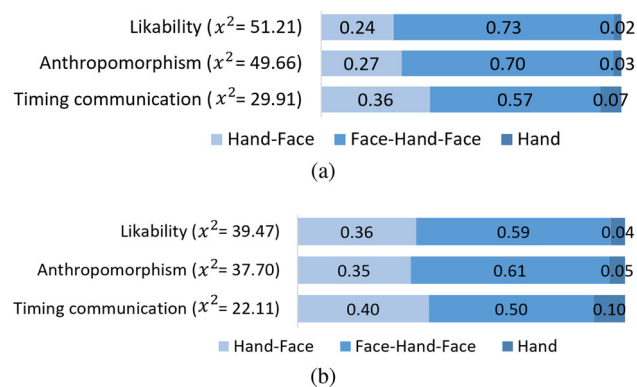

Fig. 9 χ^2 values and win-probabilities of gaze conditions in the in-person study for the three dependent measures: **a** Non-Fragile object, **b** Fragile object

5.3 Results

5.3.1 Quantitative Results

There was no bias towards selecting the first or the second handover [like: z =-0.88, p = 0.38; friendly: z = -0.27, p = 0.79; natural: z =-0.48, p = 0.63; humanlike: z = - 1.16, p = 0.25; timing communication: z =0.34, p = 0.73]. Tables 6-8 (Appendix) and Fig. 8- 10 show the robot gaze preferences of the participants in terms of likability, anthropomorphism and timing communication. In all six experimental conditions, the gaze conditions differ significantly in ratings ($p < 0.0001$), supporting **H1**. As in the video study, participants preferred the *Face-Hand-Face transition* gazes over *Hand-Face* and *Hand* gazes. *Hand* gaze was the least preferred ($p < 0.0001$).

Based on the binary proportion difference test, the proportion of participants in a handover preferring one gaze condition over other can not be claimed to be affected by object size (Table 9, Appendix), object fragility (Table 10, Appendix) and user's posture (Table 11, Appendix), con-
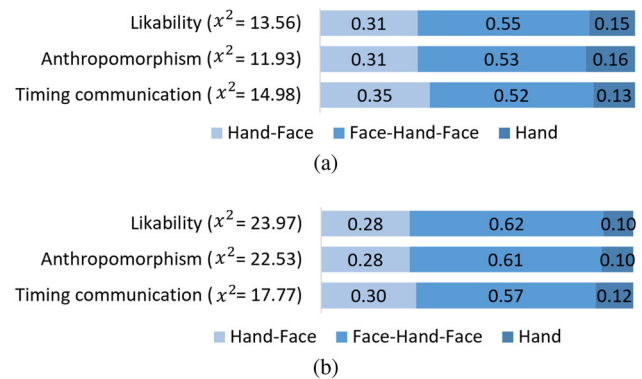

Fig. 10 χ^2 values and win-probabilities of gaze conditions in the in-person study for the three dependent measures: **a** Standing, **b** Sitting

tradicting hypotheses **H2**, **H3** and **H4**. The proportion of participants in a handover preferring one gaze condition over other (Table 12, Appendix) also cannot be claimed to be affected by the interaction modality (video or in-person), contradicting **H5**.

However, based on the equivalence tests, we did not find evidence that the proportion of participants in a handover preferring one gaze condition over the other is equivalent for the two object sizes (Table 9, Appendix), object fragilities (Table 10, Appendix), or user's postures (Table 11, Appendix). Thus, hypotheses **H2**, **H3** and **H4** can also not be rejected (all p values are over 0.15). We also did not find evidence that the proportion of participants in a handover preferring one gaze condition over other (Table 12, Appendix) is equivalent for the two interaction modalities (video or in-person). Thus hypothesis **H5** can also not be rejected.

5.3.2 Open-Ended Responses

14 out of 72 participants gave additional comments.

Seven participants made *Hand-Face* gaze vs. *Face-Hand-Face* gaze comparisons. Two of these participants stated that they preferred *Face-Hand-Face* over *Hand-Face* gaze because they preferred longer eye contact by the robot.

> P020 - *"I preferred handover 1 (Face-Hand-Face gaze) because the robot stared at me before and after the handover, and I felt accompanied by it during the entire handover."*

Four participants mentioned that they could not distinguish between the two conditions, while one participant mentioned that *Face-Hand-Face* gaze pattern didn't feel natural.

Four out of the seven participants who commented on the comparison between *Hand-Face* gaze and *Hand* gaze, said that they preferred *Hand-Face* gaze.

P014 - *"In the first handover (Hand-Face gaze) the robot looked straight at me after the handover and seemed to be more friendly."*

P050 - *"In the first handover (Hand-Face gaze), the robot's eye movement was fully accompanied by the handover movement, and therefore it seemed more natural."*

Three participants mentioned that they could not distinguish between *Hand-Face* gaze and *Hand* gaze.

Seven out of eight participants, who commented on the comparison between *Face-Hand-Face* gaze and *Hand* gaze gazes, said that they preferred *Face-Hand-Face* gaze over *Hand* gaze because of a longer eye contact by the robot.

P014 - *"In the first handover (Hand gaze), the robot focused only on the object, and in the second handover (Face-Hand-Face gaze) it focused on me too, so it felt more natural."*

P016 - *"I preferred the second handover (Face-Hand-Face gaze) mainly because the robot looked me in the eyes at the beginning and the end."*

6 Discussion

Prior works studying robot gaze in handovers did so either for a robot as giver, or—in our own prior work on robot receiver gaze—for a small and non-fragile object, and one specific posture of the human. However, for a robot receiver, the object type or giver posture might influence preferences of robot gaze behavior. This raises the question whether the findings in the prior work generalize over variations in the handover task. In this work we investigated the effect of different object types and giver postures on preferred robot gaze behavior in a human-to-robot handover. We did not find evidence that the participants' gaze preference for a robot receiver in a handover is affected by small, large, fragile and non-fragile objects, standing or sitting postures, and the interaction modality i.e. video or in-person. However, in our study, the proportion of participants preferring one gaze condition over other is not statistically equivalent. Thus we cannot completely reject the effect of these scenarios over gaze preferences. In addition, the above-mentioned prior work [6] studied the robot receiver's gaze behaviors only in the reach phase of human-to-robot handovers. The work presented in this paper **extends the empirical evidence by studying the gaze patterns for all three phases of the handover: reach, transfer and retreat**.

As in the previous study [6], results revealed that **the most preferred gaze behavior for a robot receiver was different from the observed most frequent behavior of a human receiver**. When a person receives an object from another person, the most frequent gaze behavior is a *Hand-Face* gaze, in which the receiver looks at the giver's hand throughout the reach and transfer phases, and then at the giver's face in the retreat phase. This indicates that receivers must keep their gaze focused on the task and thus sacrifice the social benefits of the face gaze. The previous findings [6] had revealed that a robot receiver can utilize the flexibility of its perception system to incorporate a face-oriented gaze for social engagement. This finding is reinforced by our current study as the participants preferred a *Face-Hand-Face* transition gaze behavior, in which, the robot initially looked at their face, then transitioned its gaze to their hand during the reach phase, continued to look at their hand during the transfer phase, and finally transitioned its gaze back to again look at their face during the retreat phase. Open-ended responses suggested that **people preferred the robot looking at their face at the beginning and the end of the handover, and the robot's eyes following the object during the transfer phase**. This gaze behavior complemented the robot's handover motion, and thus portrayed the robot as more human-like, natural, and friendly. Another possible explanation is that the social aspects of a human receiver are implicit, whereas a robot has to establish its social agency for a better handover experience. Based on these findings, **we recommend to HRI designers to implement a *Face-Hand-Face* transition gaze when the robot receives an object from a human, regardless of human posture and characteristics of the object being handed over.**

There are several limitations of this study which could motivate future work. The results are limited by the sample size and the specific cultural and demographic makeup of its participants. Larger population samples of different age groups, backgrounds, and cultures should be investigated to help generalize the findings of our experiments. Moreover, as with any experimental study, there is a question of external validity. A handover that is part of a more complex collaborative or assistive task might elicit different expectations of the robot's gaze, a fact that should be considered by designers of HRI systems. To better understand these contextual requirements, additional realistic scenarios of assistive and collaborative tasks should be considered.

7 Conclusion

Video watching studies and in-person studies of robot gaze behaviours in human to robot handovers, revealed that:

- The participants preferred a gaze pattern in which the robot initially looks at their face and then transitions its gaze to their hand and then transitions its gaze back to look at their face again.

Table 12 Results of binary proportion difference test and equivalence test for unmatched pairs comparing video and in-person user's preferences of robot gaze in handovers. Gaze condition in bold is the preferred choice in each pairwise comparison. L: Likability, A: Anthropomorphism, T: Timing communication

	Study Scenario	Metrics	Gaze Conditions	Binary		Equivalence Test	
				Test Statistic	Pvalue	Test Statistic	Pvalue
Object Size	Small Object	L	**Hand-Face** vs. Face-Hand-Face	−0.42	0.67	0.39	0.35
			Hand-Face vs. Hand	0.11	0.91	−0.92	0.18
			Face-Hand-Face vs. Hand	1.48	0.14	0.19	0.58
		A	**Hand-Face** vs. Face-Hand-Face	−0.51	0.61	0.31	0.38
			Hand-Face vs. Hand	−0.21	0.83	0.80	0.21
			Face-Hand-Face vs. Hand	0.85	0.40	−0.32	0.38
		T	**Hand-Face** vs. Face-Hand-Face	0.00	1.00	−0.75	0.23
			Hand-Face vs. Hand	−1.23	0.22	−0.25	0.60
			Face-Hand-Face vs. Hand	0.99	0.32	0.04	0.52
	Large Object	L	**Hand-Face** vs. Face-Hand-Face	−0.49	0.62	0.29	0.38
			Hand-Face vs. Hand	0.41	0.68	−0.57	0.28
			Face-Hand-Face vs. Hand	0.11	0.91	−0.92	0.18
		A	**Hand-Face** vs. Face-Hand-Face	0.08	0.94	−0.67	0.25
			Hand-Face vs. Hand	0.57	0.57	−0.34	0.37
			Face-Hand-Face vs. Hand	−0.22	0.83	0.83	0.20
		T	**Hand-Face** vs. Face-Hand-Face	−0.79	0.43	−0.03	0.51
			Hand-Face vs. Hand	0.54	0.59	−0.33	0.37
			Face-Hand-Face vs. Hand	−0.51	0.61	0.31	0.38
Object Stiffness	Non-Fragile Object	L	**Hand-Face** vs. Face-Hand-Face	0.17	0.87	−0.63	0.26
			Hand-Face vs. Hand	−1.05	0.29	−0.04	0.52
			Face-Hand-Face vs. Hand	−1.24	0.21	−0.05	0.52
		A	**Hand-Face** vs. Face-Hand-Face	−0.08	0.94	0.71	0.24
			Hand-Face vs. Hand	−1.59	0.11	−0.65	0.74
			Face-Hand-Face vs. Hand	−1.61	0.11	−0.60	0.72
		T	**Hand-Face** vs. Face-Hand-Face	−0.31	0.76	0.44	0.33
			Hand-Face vs. Hand	−0.82	0.41	0.16	0.43
			Face-Hand-Face vs. Hand	−1.03	0.30	−0.21	0.58
	Fragile Object	L	**Hand-Face** vs. Face-Hand-Face	−0.23	0.82	0.51	0.30
			Hand-Face vs. Hand	−0.74	0.46	0.45	0.33
			Face-Hand-Face vs. Hand	−0.11	0.91	0.96	0.17
		A	**Hand-Face** vs. Face-Hand-Face	−0.15	0.88	0.59	0.28
			Hand-Face vs. Hand	−0.93	0.35	0.06	0.47
			Face-Hand-Face vs. Hand	−0.73	0.47	0.27	0.39
		T	**Hand-Face** vs. Face-Hand-Face	−0.46	0.65	0.27	0.39
			Hand-Face vs. Hand	−0.60	0.55	0.36	0.36
			Face-Hand-Face vs. Hand	−0.49	0.62	0.30	0.38
User's Posture	Standing	L	**Hand-Face** vs. Face-Hand-Face	−1.00	0.32	−0.20	0.58
			Hand-Face vs. Hand	1.27	0.20	0.41	0.66
			Face-Hand-Face vs. Hand	1.19	0.23	0.32	0.62
		A	**Hand-Face** vs. Face-Hand-Face	−0.31	0.76	0.43	0.33
			Hand-Face vs. Hand	1.37	0.17	0.56	0.71
			Face-Hand-Face vs. Hand	0.62	0.54	−0.23	0.41
		T	**Hand-Face** vs. Face-Hand-Face	−0.77	0.44	−0.03	0.51
			Hand-Face vs. Hand	0.71	0.48	−0.14	0.44

Table 12 continued

Study Scenario	Metrics	Gaze Conditions	Binary		Equivalence Test	
			Test Statistic	Pvalue	Test Statistic	Pvalue
		Face-Hand-Face vs. Hand	0.17	0.87	−0.66	0.25
Sitting	L	**Hand-Face** vs. Face-Hand-Face	0.47	0.64	−0.28	0.39
		Hand-Face vs. Hand	0.36	0.72	−0.51	0.31
		Face-Hand-Face vs. Hand	−0.19	0.85	0.72	0.24
	A	**Hand-Face** vs. Face-Hand-Face	0.53	0.60	−0.20	0.42
		Hand-Face vs. Hand	0.43	0.67	−0.39	0.35
		Face-Hand-Face vs. Hand	−0.82	0.41	0.06	0.48
	T	**Hand-Face** vs. Face-Hand-Face	0.32	0.75	−0.45	0.33
		Hand-Face vs. Hand	−0.34	0.73	0.48	0.32
		Face-Hand-Face vs. Hand	0.52	0.60	−0.32	0.38

- The participants' gaze preference did not change for changes in the object size, object fragility, or the user's posture. However, the gaze preferences were also not statistically equivalent for different object size, object fragility, or the user's posture.

These results could help the design of non-verbal cues in human-to-robot object handovers, which are integral to collaborative and assistive tasks in the workplace and at home.

Availability of data and material: Our annotations of gazes in human-human handovers are available at: https://github.com/alapkshirsagar/handover-gaze-annotations/ The videos of robot gaze conditions used in our studies are available at: https://youtu.be/9dD1YHG2Nco

Declarations

Appendix

Tables 3 - 8 show the robot gaze preferences of the participants in terms of Likability, Anthropomorphism and Timing Communication. The values in the first three columns indicate the number of "wins" of a row condition over a column condition i.e. the number of participants who preferred a row condition over a column condition. For example, in Table 3 a Likability rating of 21 in the small object, "Hand-Face" row and "Hand" column shows that 21 participants liked the *Hand-Face* gaze over the *Hand* gaze. We obtained these ratings by averaging the participants' responses for both ordered pairwise comparisons, and thus some of these values are fractions. The values in a_i column show the sum of the ratings for each row. The probability that a row condition is preferred over other conditions was calculated using an iterative estimation algorithm [18] and the probability values are shown in P_i column.

Tables 9-11 show the results of binary proportion difference tests and equivalence tests for matched pairs which we used to evaluate **H2-H4**. We evaluated the user's preferred gaze behavior in terms of Likability, Anthropomorphism, and Timing Communication for different study conditions. The values in "Z-score" column represent the test statistic. For example, in Table 9, a Z-score of 0.00 and a P-value of 0.5 for Likability in *Hand-Face* vs. *Hand* gaze conditions means that the proportion of participants in the video study who liked *Hand-Face* over *Hand* condition for both small and large object is not statistically different. However, for the same scenario, a Z-score of -0.92 and a P-value of 0.18 for the Equivalence Test indicates that the proportions are not statistically equivalent as well.

Table 12 show the results of binary proportion difference tests for unmatched pairs which we used to evaluate **H5**.

References

1. Kshirsagar A, Kress-Gazit H, Hoffman G (2019) Specifying and synthesizing human-robot handovers, In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Macau

2. Moon A, Troniak D, Gleeson B, Pan M, Zheng M, Blumer B, MacLean K, Croft E (2014) Meet me where I'm gazing: How shared attention gaze affects human-robot handover timing, In ACM/IEEE International Conference on Human-robot Interaction (HRI). Bielefeld, Germany

3. Zheng M, Moon A, Croft E, Meng M (2015) Impacts of robot head gaze on robot-to-human handovers. Int J Soc Robot 7(5):783–798

4. Fischer K, Jensen L, Kirstein F, Stabinger S, Erkent Ö, Shukla D, Piater J (2015) The effects of social gaze in human-robot collaborative assembly, In International Conference on Social Robotics (ICSR). France, Paris

5. Kühnlenz B, Wang Z.-Q, Kühnlenz K (2017) Impact of continuous eye contact of a humanoid robot on user experience and interactions with professional user background, In IEEE International Symposium on Robot and Human Interactive Communication (Ro-Man), Lisbon, Portugal

6. Kshirsagar A, Lim M, Christian S, Hoffman G (2020) Robot gaze behaviors in human-to-robot handovers. IEEE Robot Autom Lett 5(4):6552–6558

7. Koay KL, Sisbot EA, Syrdal DS, Walters ML, Dautenhahn K, Alami R (2007) Exploratory study of a robot approaching a person in the context of handing over an object, In AAAI Spring Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics. Stanford CA

8. Edsinger A, Kemp CC (2007) Human-robot interaction for cooperative manipulation: Handing objects to one another, In IEEE International Symposium on Robot and Human Interactive Communication (Ro-Man). Jeju, South Korea

9. Cakmak M, Srinivasa SS, Lee MK, Forlizzi J, Kiesler S (2011) Human preferences for robot-human hand-over configurations, In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). San Francisco

10. Cakmak M, Srinivasa S. S, Lee M. K, Kiesler S, Forlizzi J (2011)Using spatial and temporal contrast for fluent robot-human hand-overs, In 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)

11. zu Borgsen S. M, Bernotat J, Wachsmuth S, (2017) Hand in hand with robots: differences between experienced and naive users in human-robot handover scenarios, In International Conference on Social Robotics (ICSR). Tsukuba, Japan

12. Pan M. K, Croft E. A, Niemeyer G (2018) Evaluating social perception of human-to-robot handovers using the robot social attributes scale (rosas), In ACM/IEEE International Conference on Human-Robot Interaction (HRI), Chicago, USA

13. Flanagan J, Johansson R (2003) Action plans used in action observation. Nat 424(6950):769

14. Strabala K, Lee M, Dragan A, Forlizzi J, Srinivasa S (2012) Learning the communication of intent prior to physical collaboration, In IEEE International Symposium on Robot and Human Interactive Communication (Ro-Man). France, Paris

15. Admoni H, Dragan A, Srinivasa S, Scassellati B (2014) Deliberate delays during robot-to-human handovers improve compliance with gaze communication, In ACM/IEEE International Conference on Human-Robot Interaction (HRI). Bielefeld, Germany

16. Carfí A, Foglino F, Bruno B, Mastrogiovanni F (2019) A multi-sensor dataset of human-human handover. Data Brief 22:109–117

17. Faibish T (2021) Human-robot handovers: Human preferences and robot learning, Master's thesis, Department of Industrial Engineering and Management, Ben-Gurion university of the Negev, Beer Sheva, Israel

18. Hunter D (2003) MM algorithms for generalized bradley-terry models. Annals Statist 32(1):384–406

19. Bradley R, Terry M (1952) Rank analysis of incomplete block designs: I. the method of paired comparisons. Biometrika 39(3/4):324–345

20. Yamaoka F, Kanda T, Ishiguro H, Hagita N (2006) How contingent should a communication robot be? In ACM SIGCHI/SIGART Conference on Human-robot Interaction (HRI). Salt Lake City, USA

21. May WL, Johnson WD (1997) The validity and power of tests for equality of two correlated proportions. Statist Med 16(10):1081–1096

22. McNemar Q (1947) Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12(2):153–157

23. Morikawa T, Yanagawa T, Endou A, Yoshimura I (1996) Equivalence tests for pair-matched binary data. Bull Inf Cybern 28(1):31–46

24. Illowsky B, Dean S (2018) Introductory statistics, pp 579–584

25. Lakens D, Scheel AM, Isager PM (2018) Equivalence testing for psychological research: a tutorial. Adv Methods Pract Psychol Sci 1(2):259–269

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.