Robot Gaze Behaviors in Human-to-Robot Handovers

Alap Kshirsagar, Melanie Lim, Shemar Christian, and Guy Hoffman

Abstract—**We present the results of two studies investigating gaze behaviors of a robot receiving an object from a human. Robot gaze is an important nonverbal behavior during human-robot handovers, yet prior work has only studied robots as givers. From a frame-by-frame video analysis of human-human handovers, we identified four receiver gaze behaviors: gazing at the giver's hand, gazing at their face, and two kinds of face-hand transition gazes. We implemented these behaviors on a robot arm equipped with an anthropomorphic head. In Study 1, participants compared videos of a handover from a human actor to a robot exhibiting these four gaze behaviors. We found that when the robot transitions its head gaze from the giver's face to the giver's hand, participants consider the handover to be more likable, anthropomorphic, and communicative of timing ($p < 0.0001$). In Study 2, participants physically performed object handovers with the robot and rated their experiences of the handovers for each of the four gaze behaviors of the robot. We found weaker effects with face gaze rated the most likable ($p = 0.01$) and anthropomorphic ($p = 0.03$) behavior. In contrast to previous studies, we found no evidence that the robot's gaze affected the start time of the human's handover.**

Index Terms—**Social human-robot interaction, physical human-robot interaction, human-centered robotics.**

I. INTRODUCTION

IN THIS work, we study the effects of a robot's head gaze behavior when it is receiving an object from a human, on the perceived liking, anthropomorphism, and timing of the handover.

Industries are rapidly introducing collaborative robotic arms alongside human workers. In addition, robotic arms are projected to be used for assistive and domestic purposes [1]. Object handover is one of the most common skills required for such a collaborative or assistive robot. Tasks such as surgical assistance, housekeeping, rehabilitation assistance and collaborative assembly require a robot to give objects to a human (robot-to-human handover) and take objects from a human (human-to-robot handover). This seemingly simple action involves coordination in both time and space of hand movements, grip forces, body postures, and other nonverbal cues like head and eye gaze. Therefore, researchers have studied human-human object handovers to understand how people perform this complex maneuver [2]–[4]. The insights from studying human-human handovers have been used for developing controllers for human-robot handovers [5]–[7].

Past research has shown that the robot's head gaze behaviors affect the subjective experience and timing of handovers [8]–[11]. However, previous works only analyzed robot-to-human handovers, in which the robot was the giver. To the best of our knowledge, no work has studied the robot's head gaze behaviors in human-to-robot handovers, in which the robot is the receiver. This lack of literature is surprising, given that human-to-robot handovers are as common and as complex a problem in collaborative and assistive robotics. We seek to address the gap by first analyzing a dataset of human-human handovers and identifying the most common gaze behaviors performed by receivers. We then implement these gaze behaviors on a robot and conduct two human-to-robot handover studies, a video study and an in-person study, to compare people's experiences of handovers with these gaze behaviors. Our contributions in this work are:

1) Evidence that—for observers of a handover—a transition gaze, in which the robot initially looks at the giver's face and then at the giver's hand, is perceived as more anthropomorphic, likable and communicative of timing compared to continuously looking at the giver's face or hand.
2) Weaker evidence that—for participants in a handover—continuously looking at the giver's face or initially looking at the face and then at the giver's hand is perceived as more anthropomorphic and likable compared to continuously looking at the giver's hand.
3) No evidence that a robot's head gaze behaviors have an effect on the reach-start time of a giver, defined as the difference between the starting of the robot's and the human's motions, in robot initiated human-to-robot handovers.
4) An additional contribution of this work is the publication of a categorized dataset of frame-by-frame head gaze labels extracted from a public dataset of handovers [12]. This complements previous studies on giver's gaze behaviors in human-human handovers [8] and provides new insights on the receiver's gaze behaviors.

Manuscript received February 24, 2020; accepted July 22, 2020. Date of publication August 11, 2020; date of current version August 20, 2020. This letter was recommended for publication by Associate Editor H. S. Ahn and Editor D. Lee upon evaluation of the reviewers' comments. *(Corresponding author: Alap Kshirsagar.)*

Alap Kshirsagar, Shemar Christian, and Guy Hoffman are with the Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY 14850 USA (e-mail: ak2458@cornell.edu; stc77@cornell.edu; hoffman@cornell.edu).

Melanie Lim is with the Systems Engineering Program, Cornell University, Ithaca, NY 14850 USA (e-mail: ml876@cornell.edu).

This letter has supplementary downloadable material available at https://ieeexplore.ieee.org, provided by the authors.

Digital Object Identifier 10.1109/LRA.2020.3015692

II. RELATED WORK

We position this work in the context of gaze behaviors in human-human handovers, as well as in the literature on human-robot handovers.

A. Gaze in Human-Human Handovers

Surprisingly, gaze behaviors in human-to-human handovers have rarely been studied. Flanagan *et al.* [13] identified that during an object passing task, human eye movements and gazes focused on the end location where objects were placed rather than the grasped object and hand in motion. In the case of a collaborative task, Macdonald *et al.* [14] observed that humans more often gazed at the shared task than at each others' faces. Other evidence comes from the human-robot interaction (HRI) literature: Strabala *et al.* [15] studied a physical collaboration task involving handovers and found that people use gaze cues to signal intent to start a handover. Shi *et al.* [16], [17] investigated human-human handovers in the context of flyer distribution to pedestrians. They found that the givers gazed at the receiver throughout the handover to maintain eye contact. Moon *et al.* [8] also studied a human giver's gaze in human-to-human handovers and found object gaze to be the most frequent gaze cues during the handover. All of the above human-human studies performed in the context of HRI only examined the giver's gaze.

B. Gaze in Human-Robot Handovers

In HRI, robot gaze has been studied widely as a nonverbal cue. It has been used for common ground in collaborative tasks [18], to increase engagement with the robot [19], and as a social behavior in storytelling [20] and teaching [21].

Several researchers have studied the effects of robot gaze in robot-to-human handovers [8]–[11]. Moon *et al.* [8] and Zheng *et al.* [9] demonstrated that gaze behaviors of the robot giver can affect the human receiver's experience and timing of the handover. Moon *et al.* compared three gaze behaviors: one in which the robot continuously looked at the handover location, one in which the robot shifted its gaze from the handover location to the receiver's face midway through the handover motion, and a baseline condition in which the robot kept looking at the ground. They found that when the robot exhibited the handover location gaze behavior, participants reached for the object earlier than the other conditions. Zheng *et al.* [9] extended this work and found that when the robot continuously gazed at the human's face, participants reached for the object even earlier compared to the handover location gaze. Also, when the robot looked at the person's face, either continuously or for some amount of time, the participants considered the handover to be more likable and anthropomorphic.

Fischer *et al.* [10] compared two gaze behaviors of a robot tasked with retrieving parts and found that when the robot looked at the person's face instead of looking at the movement of its own arm, participants engaged more with the robot and felt more responsible for the task. Kühnlenz *et al.* [11] compared gaze behaviors of a humanoid robot in a fetch-and-give scenario with the robot looking either towards the planned path or towards the human's face while approaching and handing over the object to the human. They found that looking at the human's face increased the anthropomorphism, animacy, perceived intelligence, and social presence of the robot. This literature suggests that while humans mostly perform task-oriented gazes toward the hand, object, or workspace, a robot's eye contact gaze may improve likability, anthropomorphism, and engagement.

We present a study similar to Zheng *et al.* [9], but contrary to all prior studies, we examine gaze behaviors when the robot is the receiver and base our robot behaviors on the analysis of a large dataset of human-human handovers. Human-to-robot handovers are as challenging as robot-to-human handovers in terms of control and interaction. Functionally, the robot must predict the handover location and timing. The robot also has to communicate its intent, readiness and involvement in the interaction. Robot gaze serves these functions differently in the robot-as-receiver scenario than in the robot-as-giver scenario and may affect the sequence and timing of the human's giving actions, as well as their perception of the robot. While there is some work done on predicting the handover location in human-to-robot handovers [22], [23], there is no prior work studying robot gaze communication in this scenario. Our results could be useful for designing gaze behaviors of a collaborative robot when it takes on the role of a receiver.

III. ANALYSIS OF GAZE IN HUMAN-HUMAN HANDOVERS

To find the most frequent gaze behaviors in a handover, we first performed a frame-by-frame video analysis of a public dataset of human-human handovers [12] (Fig. 1). The dataset consists of more than 1000 videos of object handovers with 18 volunteers, 10 objects, and several handover scenarios. The handover scenarios vary in terms of experiment type (volunteer-volunteer or volunteer-experimenter), role of the volunteer (giver or receiver), and starting phase (with approach or without approach). We only considered the volunteer-volunteer handovers as these would be more natural. This gave us a total of 288 videos each for givers and receivers recorded at 8fps and a resolution of 1280×720 pixels. In total we coded 14214 frames of handover videos.

We annotated 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) **R's gaze**: G's face, G's hand, Own hand, Other

To validate our video-coding scheme, we had another coder annotate a part of the data consisting of 64 out of 288 videos, which makes for 22.2% of the total data. We found the inter-coder agreement to be 80.9%.

The remainder of this paper focuses on the gaze behavior of the receiver, but for completeness of our secondary contribution to the literature on human-human handovers we provide annotations for both the giver's and receiver's gazes.

In this study, we focused on the initial ("reach") phase of the handover [2], in which both agents extend their hands towards the handover location. Fig. 2 shows the summary of video

---

[1] The annotations are available at: [Online]. Available: https://github.com/alapkshisagar/handover-gaze-annotations/


(a) (b) (c)

Fig. 1. Types of gazes observed in the human-human handovers dataset [12]. On the left is the giver and on the right the receiver: (a) Both the giver and the receiver are gazing at the other's face, (b) The giver is gazing at the other's hand while the receiver is gazing at the other's face, (c) Both the giver and receiver are gazing at the other's hand.

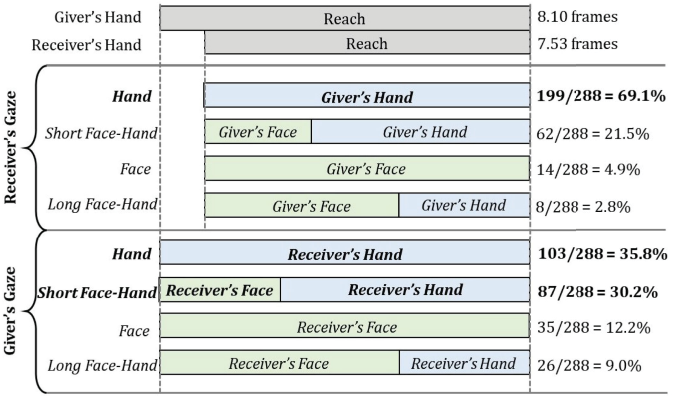

Fig. 2. Analysis of gaze behaviors in the reach phase of human-human handovers: The most frequent gaze behavior in the reach phase of the handover is *Hand* gaze in which the person continuously looks at the other person's hand.

annotations for this phase. Analysis of the annotations revealed the four most common gaze patterns as:

1) *Hand gaze:* Continuously look at the other person's hand.
2) *Face gaze:* Continuously look at the other person's face.
3) *Short Face-Hand gaze:* Initially look at the other person's face and then at their hand. The duration of the *Face* gaze is less than or equal to half the duration of the reach phase.
4) *Long Face-Hand gaze:* Initially look at the other person's face and then at their hand. The duration of *Face* gaze is more than half of the duration of the reach phase.

For both givers and receivers, gazing at the other participant's hand was the most frequent behavior. However, this was more pronounced for the receiver's gaze, with almost double the frequency (69.1% vs. 35.8%) for the *Hand* gaze. This initial finding emphasizes the need for a separate study of robot handover gaze in the robot-as-receiver role.

IV. HUMAN-ROBOT HANDOVER STUDIES

For designing gaze behaviors that a robot should exhibit when it receives an object from a human, our analysis of human-human handovers provided four candidate gaze patterns. We conducted two within subject studies, a video study and an in-person study, to investigate the likeability, anthropomorphism and timing communication of these gaze behaviors. Both studies were approved by an Institutional Review Board (IRB).

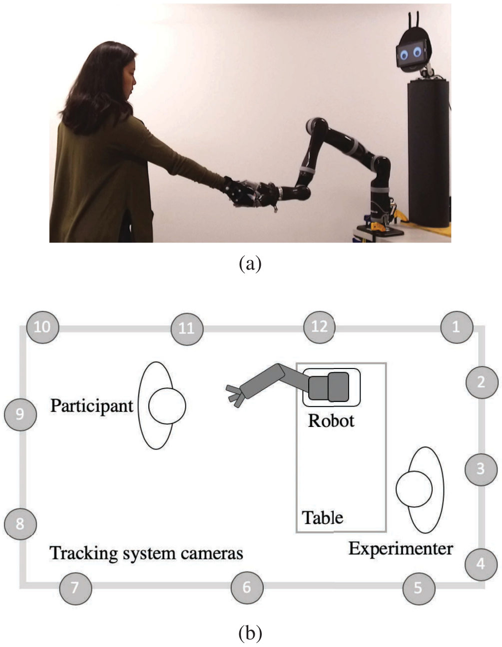

Fig. 3. The experiment setup consisted of a Kinova Jaco-2 robot arm, a robot head and an OptiTrack motion tracking system with 12 cameras. (a) shows a video frame of an actor handing over an object to the robot, used in the video study. (b) shows a diagram of the setup for the in-person study.

A. *Video Study of Human-Robot Handovers*

In the first study, participants watched and compared the robot gaze behaviors in videos of human-to-robot handovers. We recorded videos of an actor handing over an object to a Kinova Jaco-2 robot arm, accompanied by a simple robot head. The video attachment shows the recordings of robot gaze behaviours. Fig. 3(a) shows a snapshot of a video recording. The robot arm had seven degrees-of-freedom and a three-fingered gripper. The robot head had four degrees-of-freedom and a 7-inch screen which displayed a static image of two eyes. We used an OptiTrack motion tracking system to track the positions of the human's hand, the robot gripper and the object. The robot arm was autonomous and programmed to reach a predefined

TABLE I
COMBINED PREFERENCES OF GAZE BEHAVIORS IN THE VIDEO STUDY. LARGER a_i AND P_i INDICATE A STRONGER PREFERENCE TO THE ROW CONDITION. L-FH = *LONG FACE-HAND TRANSITION* GAZE, S-FH = *SHORT FACE-HAND TRANSITION* GAZE

		Face	Hand	L-FH	S-FH	a_i	P_i
Likability	Face	0	6.75	2.75	2.25	11.75	0.04
	Hand	17.25	0	3.5	4.75	25.5	0.09
	L-FH	21.25	20.5	0	8	49.75	0.34
	S-FH	21.75	19.25	16	0	57	0.52
Anthropomorphism	Face	0	5	3.5	3.25	11.75	0.05
	Hand	19	0	5	4.75	28.75	0.12
	L-FH	20.5	19	0	8.5	48	0.33
	S-FH	20.75	19.25	15.5	0	55.5	0.50
Timing Communication	Face	0	5.5	3	2	10.5	0.03
	Hand	18.5	0	3.5	4	26	0.09
	L-FH	21	20.5	0	7	48.5	0.31
	S-FH	22	20	17	0	59	0.57

position once the handover began. The robot grasped the object when the object was close enough. Finally, the robot retreated to its home position after the human released the object and started to retreat.

We programmed the head of the robot to exhibit the following four gaze behaviors:

i) *Face gaze:* Continuously look at the giver's face. The face gaze location was programmed manually to a fixed location.
ii) *Hand gaze:* Continuously look at the giver's hand.
iii) *Short Face-Hand transition gaze:* Initially look at the giver's face and transition to the giver's hand. The transition began as soon as the reach phase started. Thus the *Face* gaze was shorter than the *Hand* gaze.
iv) *Long Face-Hand transition gaze:* Initially look at the giver's face and transition to the giver's hand. The transition began as soon as the robot's reach phase ended. Thus the *Face* gaze was longer than the *Hand* gaze.

1) Procedure: The study was conducted in a laboratory environment and the experimenter left the room after the participant started the study. Participants gave online consent, read the instructions and then completed a practice session followed by 12 actual study sessions. In each session they watched two handover videos, one after the other. The 12 sessions consisted of the six possible pairings of the four gaze patterns and their reverse order. 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 head movement in the video. After every two videos, they were asked to answer the following questions:

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)

This questionnaire is identical to the one in Zheng *et al.* [9]. 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*.

2) Participants: A total of 24 participants participated in the experiment (13 Male, 11 Female). Participants were recruited through emails and posters. Each study session lasted for about 15 minutes. The participants were compensated with a $5 gift card for participating in the study. The order of the videos was randomized and counterbalanced.

3) Preference Ranking Method: Participants had evaluated the six pairings of four gaze conditions and their reverse order (for example, *Face* gaze vs. *Hand* gaze and *Hand* gaze vs. *Face* gaze). We used one-sample Wilcoxon signed-rank tests to check if participants exhibited any bias towards selecting first or the second handover. We did not find a bias [likeability: $S = -156.5$. $p = 0.90$, anthropomorphism: $S = -610, p = 0.63$, timing communication: $S = 0, p = 1.00$].

Table I shows the number of participants, out of the 24 participants, who chose the row condition over the column condition. For example, a rating of 5 for *anthropomorphism* in row "Face" and column "Hand" indicates that five participants chose *Face* gaze over *Hand* gaze in the pairwise comparisons. a_i represents the number of times row condition i "wins" against other conditions.[2] P_i is the probability that row condition i is preferred over other conditions. We used the iterative estimation algorithm proposed by Hunter [24] to compute the P_i values.

4) Hypotheses: Both studies tested the same single hypothesis across all three dependent measures, namely that there is a difference in the probability for preference as a consequence of the gaze type, i.e. $P_i \neq P_j \; \forall i \neq j$. We did not have a-priori hypotheses about the order of the conditions.

5) Quantitative Results: We used the Bradley-Terry method [25] to evaluate participants' rankings of the likeability, anthropomorphism and timing communication of gaze behaviors.[3] The results are shown in Fig. 4. As all of the χ^2 values are large ($p < 0.0001$), we conclude that the gaze

---

[2] The ratings for each row were obtained by averaging the ratings for both ordered pairwise comparisons. a_i shows the sum of ratings in each row. Since each subject compared condition i with $g - 1$ conditions, where $g = 4$ is the number of gaze behaviors, a_i represents the number of "wins" against other conditions in $n \times (g - 1)$ comparisons, where n is the number of participants.

[3] Following the procedure used by Yamaoka *et. al* [26] for each metric, the value of $\chi_0^2 = ng(g-1)ln2 - 2Bln10$, where $B = n\sum_{i<j} log(P_i + P_j) - \sum_i a_i log P_i$, should be greater than the $\alpha = 0.005$ point of a χ^2 distribution with $(g - 1)$ degrees of freedom.

Fig. 4. χ^2 values and win-probabilities of gaze conditions in the video study for the three dependent measures. (For *Face* gaze, the win-probabilities are 0.04, 0.05, 0.03 for Likability, Anthropomorphism and Timing Communication, respectively.)

condition affects ratings. Participants prefer the *Face-Hand transition* gazes more than *Hand* and *Face* gazes. Also, *Face* gaze is the least preferred condition.

6) Open-Ended Responses: Participants were asked to write optional comments after each session. 16 out of 24 participants gave at least one additional comment.

Eight participants made *Face* vs *Hand* gaze comparisons. Four said that they preferred *Hand* over *Face* gaze because it signalled attentiveness, while three said they preferred *Face* gaze because *Hand* gaze signalled shyness:

> P010: *"The aparent attentiveness in the second [Hand gaze] was reassurance [sic]."*

> P023: *"I felt that in the 2nd one [Hand gaze] it was shy."*

Seven participants implied at least once that they preferred some head movement over no head movement.

> P007: *"Some movement of the head is more friendly than no movement."*

> P006: *"The lack of a head tilt in the second one seems much more artificial and less human-like."*

Four out of ten participants, who commented on comparing the two transition gazes, mentioned that they could not distinguish between them, while five of them said that they preferred *Short Face-Hand* transition gaze.

> P007: *"Both [transition gazes] seemed identical."*

> P006: *"The initiation of the head tilt in the second one [Short Face-Hand transition gaze] made it more apparent when the robot wanted to receive the object, first one tilted the head after it was already being handed the object."*

B. In-Person Study of Human-to-Robot Handovers

To study people's perceptions of robot gaze behaviors in handovers with a real robot, we ran a second study. A separate set of participants performed object handovers with the Kinova Jaco-2 arm, accompanied by the same simple robot head. The setup is shown in Fig. 3 (b). The robot arm and the robot head were programmed in the same way as the video study described in Section IV-A.

1) Procedure: After entering the lab, participants electronically signed the consent form and answered a question on their familiarity with a collaborative robot such as the one shown. The mean familiarity with this type of robot was found to be low (M = 1.29, SD = 0.55, on a scale of 1–5).

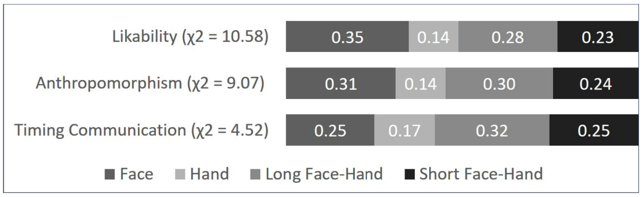

Fig. 5. χ^2 values and ratings of gaze conditions in the in-person study for the three dependent measures.

The experimenter then verbally described the experiment and participants completed a practice session. This was followed by 12 study sessions. In each session, the participants performed two handovers with the robot, sequentially. The object being handed over was a partially filled water bottle weighing about 200 g. The 12 sessions consisted of the six possible pairings of the four gaze patterns and their reverse order. After every session they were asked to answer the same choice questions as in the video study.

2) Participants: A total of 24 participants, different from the participants in the video study, participated in the experiments (6 Male, 18 Female). Participants were recruited through an online recruitment system. Each study session lasted for about 15 minutes. The participants were compensated with a $5 gift card. The order of the conditions was randomized and counterbalanced.

3) Metrics and Hypotheses: We used the same preference ranking method as in Study 1, described in Section IV-A3.

Table II shows the number of participants, out of the 24 participants, who chose the row condition over the column condition. We had the same three hypotheses about overall differences in preferences caused by the gaze manipulation.

In addition, in Study 2, we measured the human's *reach start time* in each handover, defined as the difference between the start of the human hand's reaching motion and the start of the robot arm's reaching motion. Previous studies on robot-to-human handovers found that the robot's gaze behavior affected the human's reach start time [8], [9]. We were interested to see if the robot's gaze behavior had a similar effect in a human-to-robot handover.

4) Quantitative Results: We used one-sample Wilcoxon signed-rank tests to check if participants exhibited any bias towards selecting the first or the second handover. We did not find a bias [likeability: $S = -1481$, $p = 0.25$, anthropomorphism: $S = -684$, $p = 0.59$, timing communication: $S = -289$, $p = 0.81$]. We conducted a one-way ANOVA on the human's *reach start time* for four gaze conditions. The results indicated no detectable difference between the four conditions ($F = 0.429$, $p = 0.732$).

We used the Bradley-Terry method [25] to evaluate participants' rankings of the likeability, anthropomorphism and timing communication of gaze behaviors. The results are shown in Fig. 5. Compared to Study 1, we detected a weaker effect of condition on likability and anthropomorphism, with $p = 0.01$ and $p = 0.03$ respectively, and virtually no effect on timing communication ($p = 0.21$). Participants rated all dependent measures lower in the *Hand* gaze condition than in other conditions.

5) Open-Ended Responses: 20 out of 24 participants wrote at least one additional comment. In their comments, contrary to

TABLE II

COMBINED PREFERENCES OF GAZE BEHAVIORS IN THE IN-PERSON STUDY. LARGER a_i AND P_i INDICATE A STRONGER PREFERENCE TO THE ROW CONDITION. L-FH = *LONG FACE-HAND TRANSITION* GAZE, S-FH = *SHORT FACE-HAND TRANSITION* GAZE

		Face	Hand	LongFace-Hand	ShortFace-Hand	a_i	P_i
Likability	Face	0	15	14.5	15.75	45.25	0.35
	Hand	9	0	7	7.75	23.75	0.14
	LongFace-Hand	9.5	17	0	13	39.5	0.28
	ShortFace-Hand	8.25	16.25	11	0	35.5	0.23
Anthropomorphism	Face	0	14.25	13	15	42.25	0.31
	Hand	9.75	0	6.75	7.5	24	0.14
	LongFace-Hand	11	17.25	0	13.25	41.5	0.30
	ShortFace-Hand	9	16.5	10.75	0	36.25	0.24
Timing Communication	Face	0	11.5	11.5	13.5	36.5	0.25
	Hand	12.5	0	7	8.5	28	0.17
	LongFace-Hand	12.5	17	0	13	42.5	0.32
	ShortFace-Hand	10.5	15.5	11	0	37	0.25

the comments in our video study, nine out of 14 participants said that they preferred Face gaze over the Hand gaze, as it was more engaging, human-like and friendly.

> P008 - *"With a task this simple, maintaining eye contact feels more engaged/natural than looking at the object."*

> P016 - *"In the second one [Face gaze], the robot looked me in the eye and seemed more humanlike."*

Six participants implied at least once that they preferred some head movement over no movement.

> P014 - *"No movement at all with 2nd one [Hand gaze] made it less humanlike."*

> P023 - *"The ones where the head moves seem more natural and humanlike."*

Six out of the 12 participants who gave comments on the comparison between two transition gazes said that they could not distinguish between the two transition gazes, while four said that they preferred the *Long Face-Hand transition* gaze.

> P008 - *"Both tries felt roughly the same to me."*

> P016 - *"The robot head appeared to move more naturally in the second handover [Long Face-Hand transition gaze]."*

V. DISCUSSION

Participants in a handover combine two types of gaze behavior: a task-oriented gaze toward the hand or the object, and a face-oriented gaze for social engagement. Prior studies with robots as givers showed that *Face* gaze is beneficial as it causes more positive evaluations of the robot. Unfortunately, there is not much prior literature to draw on with respect to receiver gaze in handovers, whether the receiver is a human or a robot. In our analysis of human-human handovers, most receivers used *Hand* gaze over other types of gazes, at almost double the rate as givers. This might indicate a stronger need for receivers to keep their gaze focused on the task, sacrificing the social benefits of eye contact. When the receiver is a robot, our studies paint a different picture: Task-oriented *Hand* gaze was consistently ranked as less preferred, especially when compared to a behavior that shifted from *Face* gaze to *Hand* gaze.

What about consistent face gazing? Comparing our video and in-person studies reveals different results for the preference of the *Face* gaze. When participants watched recorded videos of a human handing over an object to a robot, they preferred the *Hand* gaze over the *Face* gaze. Open ended responses suggested that looking at the hand signalled that the robot was more attentive. In contrast, when the participants physically performed the object handover with the robot, the *Face* gaze dominated the *Hand* gaze. Open ended responses suggested that looking at the face was a friendly gesture. This is in line with previous research that shows that people have different preferences when they are interacting with a robot in first person versus when they are observing video interactions of a third person with the same robot [27]. However, these previous studies showed mostly a change in degree, not a categorical shift as found here.

In the video study, participants preferred the *Face-Hand transition* gaze behaviors to consistent *Hand* or *Face* gaze behaviors, citing concerns that the lack of movement was unnatural. Despite similar concerns in the in-person study, participants' preference ratings for *Face-Hand transition* gaze and *Face* gaze were similar. All of them were higher than the *Hand* gaze, indicating that the social benefits of the *Face* gaze—especially in terms of likability—overcome some of the concerns associated with lack of movement.

Combining the findings of the two studies, our recommendation to HRI designers is to implement a *Face-Hand transition* gaze when a robot is receiving objects from a human. There seems to be no preference for the timing of the transition. The least recommended behavior is a static *Hand* gaze, even though it is the most common behavior of a human receiver.

A human receiver may have to focus on the receiving task more than a robot receiver who can be equipped with a more flexible sensor system. The robot should take advantage of this flexibility and incorporate the beneficial *Face* gaze behavior. If the robot does not have an actuated gaze, a *Face* gaze is preferred over a *Hand* gaze.

In previous studies [8], [9], researchers found that the robot's gaze behavior had an effect on the reach start time of the human in robot-to-human handovers. However, we did not find a similar effect of the robot's gaze behavior on the human's reach start time in human-to-robot handovers. A possible reason is that the repetitive handovers caused learning or fatigue, affecting the reaching behavior. However, the same prior works that found timing differences across gaze conditions also used repeated trials [8], [9]. This may suggest that, while robot gaze could serve as a timing cue in robot-to-human handovers, it might not be an effective timing cue in human-to-robot handovers.

In contrast to prior works, which conducted in-person studies, we conducted a video study alongside an in-person study. Video studies offer more manipulation control and thus more internal validity, whereas in-person studies have more ecological validity. We used a different set of participants to prevent order effects. We did not have specific hypotheses that warranted a quantitative comparison of the data from the two studies, and present the difference only qualitatively.

Our studies have limitations which could motivate future work. We chose to use head gaze over eye gaze as head movement is more prominent and noticeable even in the peripheral vision. We only considered a standing face-to-face handover without the approach phase. Other giver-receiver configurations such as standing-sitting or sitting-sitting could also be considered, or the effects of robot gaze during the approach phase of human-to-robot handovers could be investigated. In the future, we plan to investigate whether the size or shape of the object or the gender of the giver affects preferences towards gaze behaviors.

VI. CONCLUSION

We conducted a video watching study and an in-person study to investigate gaze behaviors of a robot in human-to-robot handovers. While receiving the object from the human, the robot either continuously looked at the human's hand, continuously looked at the human's face, or transitioned its gaze from the face to the hand. These gaze patterns were inspired by those exhibited by receivers in human-to-human handovers. In the video study, we found that participants liked the transition gaze behaviors more than any of the fixed gaze behaviors. They also considered the transition gazes to be more anthropomorphic. In the in-person study, we found that looking at the human's hand was the least preferred behavior, even though this was the most frequent behavior of receivers in human-to-human handovers. Our results could help the design of nonverbal cues in human-to-robot object handovers, which are integral to collaborative and assistive tasks in the workplace and at home.

REFERENCES

[1] P. Kellmeyer, O. Mueller, R. Feingold-Polak, and S. Levy-Tzedek, "Social robots in rehabilitation: A question of trust," *Sci. Robot.*, vol. 3, 2018, Paper eaat1587.

[2] K. Strabala *et al.*, "Toward seamless human-robot handovers," *J. Human-Robot Interact.*, vol. 2, no. 1, pp. 112–132, 2013.

[3] W. Chan, C. Parker, H. Van der Loos, and E. Croft, "Grip forces and load forces in handovers: Implications for designing human-robot handover controllers," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, 2012, pp. 9–16.

[4] W. Chan, M. Pan, E. Croft, and M. Inaba, "Characterization of handover orientations used by humans for efficient robot to human handovers," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2015, pp. 1–6.

[5] W. Chan, C. Parker, H. Van der Loos, and E. Croft, "A human-inspired object handover controller," *Int. J. Robot. Res.*, vol. 32, no. 8, pp. 971–983, 2013.

[6] C.-M. Huang, M. Cakmak, and B. Mutlu, "Adaptive coordination strategies for human-robot handovers," in *Proc. Robot.: Sci. Syst.*, vol. 11, 2015, Paper 31.

[7] A. Kshirsagar, H. Kress-Gazit, and G. Hoffman, "Specifying and synthesizing human-robot handovers," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2019, pp. 5930–5936.

[8] A. Moon *et al.*, "Meet me where I'm gazing: How shared attention gaze affects human-robot handover timing," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, 2014, pp. 334–341.

[9] M. Zheng, A. Moon, E. Croft, and M. Meng, "Impacts of robot head gaze on robot-to-human handovers," *Int. J. Social Robot.*, vol. 7, no. 5, pp. 783–798, 2015.

[10] K. Fischer *et al.*, "The effects of social gaze in human-robot collaborative assembly," in *Proc. Int. Conf. Social Robot.*, 2015, pp. 204–213.

[11] B. Kühnlenz, Z.-Q. Wang, and K. Kühnlenz, "Impact of continuous eye contact of a humanoid robot on user experience and interactions with professional user background," in *Proc. IEEE Int. Symp. Robot Human Interactive Commun.*, 2017, pp. 1037–1042.

[12] A. Carfí, F. Foglino, B. Bruno, and F. Mastrogiovanni, "A multi-sensor dataset of human-human handover," *Data Brief*, vol. 22, pp. 109–117, 2019.

[13] J. Flanagan and R. Johansson, "Action plans used in action observation," *Nature*, vol. 424, no. 6950, pp. 769–771, 2003.

[14] R. Macdonald and B. Tatler, "Gaze in a real-world social interaction: A dual eye-tracking study," *Quart. J. Exp. Psychol.*, vol. 71, no. 10, pp. 2162–2173, 2018.

[15] K. Strabala, M. Lee, A. Dragan, J. Forlizzi, and S. Srinivasa,"Learning the communication of intent prior to physical collaboration," in *Proc. IEEE Int. Symp. Robot Human Interactive Commun.*, 2012, pp. 968–973.

[16] C. Shi, M. Shiomi, C. Smith, T. Kanda, and H. Ishiguro, "A model of distributional handing interaction for a mobile robot," in *Proc. Robot.: Sci. Syst.*, Berlin, Germany, 2013, pp. 24–28.

[17] C. Shi, S. Satake, T. Kanda, and H. Ishiguro, "A robot that distributes flyers to pedestrians in a shopping mall," *Int. J. Social Robot.*, vol. 10, no. 4, pp. 421–437, Nov. 2017.

[18] G. Hoffman and C. Breazeal, "Collaboration in human-robot teams," in *Proc. AIAA Intell. Syst. Tech. Conf.*, 2004, p. 6434.

[19] C. Sidner and M. Dzikovska, "A first experiment in engagement for human-robot interaction in hosting activities," in *Proc. Adv. Natural Multimodal Dialogue Syst.*, 2005, pp. 55–76.

[20] B. Mutlu, J. Forlizzi, and J. Hodgins, "A storytelling robot: Modeling and evaluation of human-like gaze behavior," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, 2006, pp. 518–523.

[21] S. Andrist, T. Pejsa, B. Mutlu, and M. Gleicher, "Designing effective gaze mechanisms for virtual agents," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2012, pp. 705–714.

[22] S. Parastegari, B. Abbasi, E. Noohi, and M. Zefran, "Modeling human reaching phase in human-human object handover with application in robot-human handover," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 3597–3602.

[23] H. Nemlekar, D. Dutia, and Z. Li, "Object transfer point estimation for fluent human-robot handovers," in *Proc. Int. Conf. Robot. Automat.*, 2019, pp. 2627–2633.

[24] D. Hunter, "MM algorithms for generalized Bradley-Terry models," *Ann. Statist.*, vol. 32, no. 1, pp. 384–406, Feb. 2003.

[25] R. Bradley and M. Terry, "Rank analysis of incomplete block designs: I. the method of paired comparisons," *Biometrika*, vol. 39, no. 3/4, pp. 324–345, 1952.

[26] F. Yamaoka, T. Kanda, H. Ishiguro, and N. Hagita, "How contingent should a communication robot be?" in *Proc. ACM SIGCHI/SIGART Conf. Human-Robot Interact.*, 2006, pp. 313–320.

[27] M. Bretan, G. Hoffman, and G. Weinberg, "Emotionally expressive dynamic physical behaviors in robots," *Int. J. Human-Comput. Stud.*, vol. 78, pp. 1–16, 2015.