Deliberate Delays During Robot-to-Human Handovers Improve Compliance With Gaze Communication

Henny Admoni[1], Anca Dragan[2], Siddhartha S. Srinivasa[2], Brian Scassellati[1]
[1]Department of Computer Science, Yale University, New Haven, CT 06520
[2]The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213
{henny,scaz}@cs.yale.edu, {adragan,siddh}@cs.cmu.edu

ABSTRACT

As assistive robots become popular in factories and homes, there is greater need for natural, multi-channel communication during collaborative manipulation tasks. Non-verbal communication such as eye gaze can provide information without overloading more taxing channels like speech. However, certain collaborative tasks may draw attention away from these subtle communication modalities. For instance, robot-to-human handovers are primarily manual tasks, and human attention is therefore drawn to robot hands rather than to robot faces during handovers. In this paper, we show that a simple manipulation of a robot's handover behavior can significantly increase both awareness of the robot's eye gaze and compliance with that gaze. When eye gaze communication occurs during the robot's release of an object, delaying object release until the gaze is finished draws attention back to the robot's head, which increases conscious perception of the robot's communication. Furthermore, the handover delay increases peoples' compliance with the robot's communication over a non-delayed handover, even when compliance results in counterintuitive behavior.

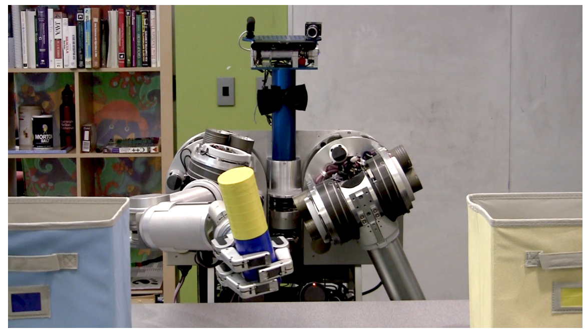
1. INTRODUCTION

In the future, assistive robots will help people perform manual tasks more easily and efficiently. These robots may retrieve items from high shelves [23], assist in fine motor manipulations [4], or act as extra hands during physically complex tasks [10]. One of the primary challenges for such robots will be the ability to manipulate objects in collaboration with people [1].

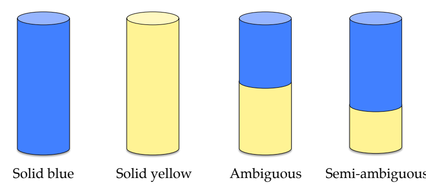
For example, imagine a robot, like the one in Figure 1a, that helps a wheelchair-bound user cook a meal. This robot can move around the kitchen, grabbing the right ingredients and handing them to the user. The robot and user can also prepare parts of the meal simultaneously, passing utensils and ingredients back and forth between them. Finally, the robot can help clean up, taking items from the user and moving them to the sink.

HRI'14, March 3–6, 2014, Bielefeld, Germany.

http://dx.doi.org/10.1145/2559636.2559682.



(a) A participant's view of the block sorting task.



(b) Blocks were fully colored, ambiguously colored (50% of each color) or semi-ambiguously colored (70% of one color).

Figure 1: Participants engage in a collaborative manipulation task with HERB. The robot hands over colored blocks, and participants sort them into colored boxes on the table.

A common task throughout this interaction is the *handover*: the act of transferring an item from one actor to another. For seamless robot-to-human handovers, the robot must generate appropriate social cues that alert the person to the what, when, and where of the handover [7, 28].

But other information, unrelated to the handover itself, may also need to be communicated during a handover. For example, the robot might want to indicate where to put an object after giving it. Speech is an obvious mode of communication for conveying such information, but it may not be available or effective in all situations. For instance, speech may be unavailable in a noisy room, when interacting with the hearing impaired, or when a person is already engaged in a listening task, such as holding a conversation while cooking. Even when speech is available, it is not always the most effective means of communicating: a robot that announced every handover before it occurred would hinder the fluency of interactions involving frequent handovers.

Eye gaze is an alternative means of communication that can be used when speech isn't practical. Typical humans can understand the motor intentions of others based on their gaze [24], and robots are able to influence human motor behavior using gaze (e.g., [22]). Based on this, a natural conclusion is to communicate information about where to put the object using eye gaze.

Because eye gaze requires the user to attend to the cue in order to be effective, it becomes critical to select the right time to exhibit the eye gaze cue. This communication should occur during the *transfer phase* of the handover, which starts from the point at which the giver has finished reaching with the object toward the agreed-upon transfer location, and ends when the receiver has taken hold of the object and retracted it, signaling the end of joint activity [28]. Earlier signals specifying where to put the object may be confused with attempts to establish the what, when, and where features of the handover [28]. Signals sent after the transfer phase may be missed, because the user's attention may have already shifted to the next task location [15]. Thus the transfer phase is the ideal time to indicate temporally relevant but non-handover-related information.

However, handovers are primarily manual tasks that draw attention directly to the robot's hand and away from its eyes. Mutual eye gaze is not a necessary part of handovers [27], and because people fixate their gaze almost exclusively on areas of their environment related to the task [15], attention is often directed somewhere other than the robot's head during a handover. In our first attempts to influence human behavior using eye gaze, we found that drawing attention to the robot's face during the gaze cue was surprisingly difficult. People responded to the unfamiliar experience of robot handovers by focusing intently on the robot's hand, ignoring all non-verbal communication from the robot's head.

To address this, we introduce the idea of a *deliberate delay*—an intentional hiccup in the handover—which prompts users to shift their attention from the robot's hand to its head. In particular, we look at deliberately delaying the transfer phase of the handover by postponing the release of an object from the robot giver to the human receiver until a gaze cue has been delivered. Handovers cause user attention to be focused on the robot's hand. By manipulating the force profile during a handover—deliberately making the robot hold on to an object longer—we can draw attention to other channels like head direction.

In this paper, we present experimental evidence that adding a deliberate delay to a handover is beneficial for communicating non-handover information non-verbally, even though this delay decreases the smoothness of the handover itself. As we show below, a deliberate delay not only increases attention to the robot's head, but also increases the rate at which people comply with the robot when its gaze cue leads to counterintuitive actions.

We test the effect of a handover delay on a simplified collaborative task (Figure 1). In this task, a robot called HERB hands colored blocks to participants, who sort those blocks into one of two colored boxes according to their personal preference. A 6-axis force-torque sensor in the robot's hand identifies when the participant has grasped the block to begin the transfer phase of the handover.

When it enters the transfer phase, the robot gives sorting suggestions by looking at one of the boxes. We manipulate when this gaze cue is executed relative to the object release during transfer: in the "no delay" case, the gaze cue and release occur simultaneously; in the "delay" case, the release is delayed until the gaze cue is complete (Figure 2).

We hypothesize that:

H1 A handover delay will cause people to pay more attention to HERB's gaze communication, and that

H2 Social gaze will lead people to comply more with HERB's counterintuitive suggestion.

Our results validate H1: a deliberate delay during the transfer phase of a handover causes people to pay more attention to the robot's head and to notice the robot's nonverbal gaze suggestion more frequently. More surprisingly, the delay also causes people to comply more often with the robot's sorting suggestion, even when it is contrary to their natural behavior. This compliance effect holds even for people who explicitly notice HERB's suggestion, indicating that the delay itself, and not solely increased attention, is responsible for increased compliance. Interestingly, H2 was not supported: we found no effect of social gaze on compliance.

This work lays the foundation for a new type of robot handover. Instead of working toward seamlessness, researchers can design behaviors that leverage alternate communication channels, such as eye gaze, and introduce targeted, deliberate imperfections to improve communication and efficiency.

2. RELATED WORK

Our work draws from two areas of research in HRI: robot-to-human handovers and robot gaze communication. Though these areas have developed independently, they share considerable overlaps, for instance, using joint attention to signal a handover. Rather than surveying the broad fields individually, we highlight papers in each area that address the overlap between handovers and gaze communication.

Handovers can be divided into three distinct phases [19, 28]: the *approach*, during which a giver moves toward a receiver; the *signal*, during which giver and receiver communicate their readiness for the handover; and the *transfer*, during which the object is transferred from giver to receiver.

Handovers are a primary part of collaborative robotics, and there is strong interest in automatically generating successful robot-to-human handovers [2, 12, 25, 28]. To produce a successful handover, a robot must first convey its intention to execute that handover, which requires both spatial information (a distinct handover pose) and temporal contrast (a distinct movement profile for handovers) [7, 13]. HRI studies have attempted to determine user preferences for optimal handover behavior, such as maximal arm extension [6], minimum jerk motion profiles [17], and legibility of motion [11]. Metrics for determining human preferences range from surveys and observation [6] to physiological measurements like skin conductance and eye movement [9].

The structure of human-human handovers can also be used to inform human-robot handovers [28]. Investigations of human-human handovers identified that object transfer time (from initial contact by the receiver to final release from the giver) is approximately 500 milliseconds [8].

Mutual gaze is not a predictor of handover initiation; confirming the partner's availability through asynchronous fixations is more important to successful handovers than synchronized mutual eye contact [27]. However, taking a human partner's eye gaze into account when planning a handover increases the success of robot-to-human handovers [14]. More generally, human gaze is task driven, and gaze fixations are rarely directed to locations in the world that are not relevant to the task, even if they are visually salient [15]. Fixations are instead guided by the spatio-temporal requirements of

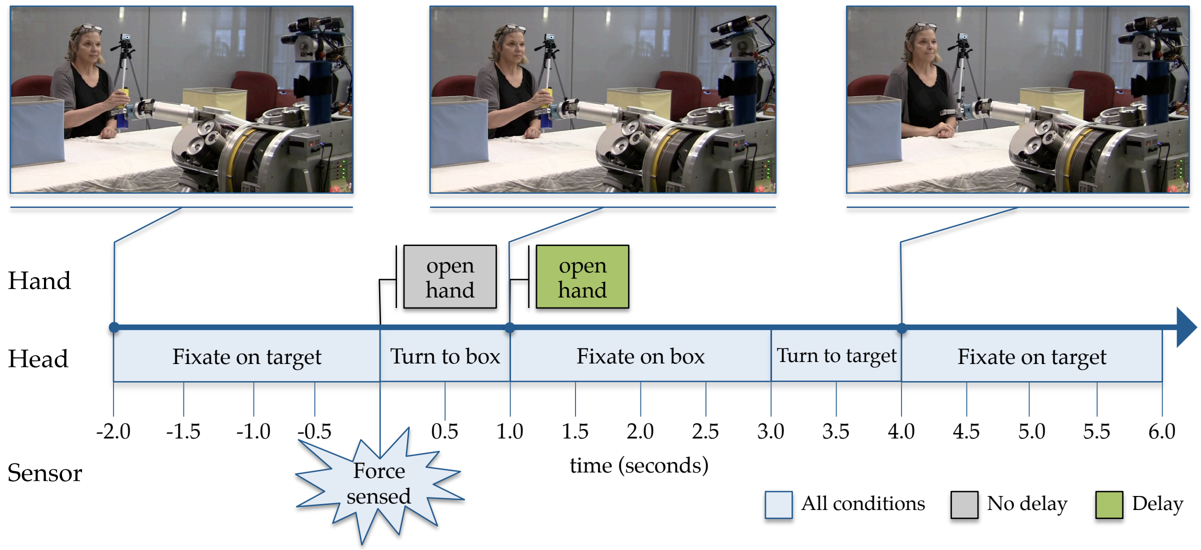


Figure 2: A graphical timeline of HERB's head and hand motions turning the transfer phase of the handover. HERB begins by fixating on a target (the participant's face or a mirrored point, depending on gaze condition), then looks at one of the boxes as a suggestion, and finally returns to fixate on the target. The timing of the block release depends on delay condition; it is either simultaneous with the head turn toward a box, or just after the head turn.

the task, arriving at the relevant location just at the point at which they are needed for task completion [18].

Robot eye gaze, however, is an important communication mode in HRI. Robots can communicate information through gaze during tasks like storytelling [20] and teaching [3]. Robot gaze cues such as joint attention facilitate performance in cooperative tasks [5] and improve perceptions of a robot's competence and naturalness [16]. Robots can even manipulate peoples' behavior using only gaze cues, prompting people to adopt certain conversational roles [21] or select certain objects from a set [22], even without people consciously registering the gaze cue. Robot gaze is clearly an informative communication channel in human-robot handovers, and we leverage this to provide sorting suggestions in the experiment described here.

3. METHODS

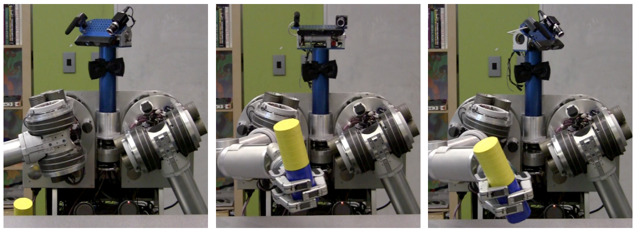
Users engage in a simple collaborative manipulation task with HERB: the robot hands blocks to participants, who are asked to sort these blocks into either a yellow box or a blue box on the table in front of them. Most blocks are unambiguously colored (fully yellow or blue), but some blocks are ambiguously colored (50% yellow and 50% blue) or semi-ambiguously colored (70% of one color and 30% of another), as illustrated in Figure 1b.

The main manipulation in this experiment is the deliberate delay between HERB's sorting suggestion and block release (Figure 2). HERB provides a non-verbal sorting suggestion by looking (i.e., orienting its head) toward one of the boxes on the table. On "no delay" trials, HERB simultaneously releases the block to the participant and executes the suggestion behavior by looking at one of the boxes. On "delay" trials, HERB first looks at one of the boxes, and only then releases the block to the participant. The suggestion behavior takes about four seconds to execute: one second for HERB to turn its head toward the box, two seconds to gaze at the box, and one second to return its head to the starting point. Thus, there is a one second difference in when the block is released between no delay and delay conditions.

There are several strategies participants could employ for sorting the blocks. One of the more obvious strategies is to sort by dominant color, putting the primarily yellow blocks in the yellow box and the primarily blue blocks in the blue box. Another strategy is to sort by the first visible color (typically the top color), regardless of how the rest of the block is colored. Other strategies, like sorting randomly or alternating boxes, do not take color into account. There are several possible sorting strategies; our analyses do not rely on participants to follow any particular strategy.

In order to test whether people see and comply with a robot's suggestions, we made HERB's suggestions as counterintuitive as possible. Therefore, when handing over the ambiguous block, HERB always suggested the bottom (i.e., less visible) color when it presented the block, which conflicts with the top color strategy (Figure 3). Because the block was exactly half of each color, however, there was no conflict with the dominant color strategy, so the ambiguous case was only mildly counterintuitive. When handing over the semi-ambiguous block, HERB always presented the block with the dominant color on top and always suggested the less dominant color. HERB's suggestion conflicts with both the dominant color strategy *and* the top color strategy, making this a highly counterintuitive suggestion.

We also manipulated whether HERB engages in social or non-social gaze before the suggestion (Figure 4). In the joint attention condition, HERB first makes eye contact by looking at the participant's face, then down at the block in its hand, and then back to the participant's face as it reaches with the block to begin the handover. After HERB initiates the suggestion by turning its head to one of the boxes, it again returns to look at the participant's face before retracting its hand. In the mirrored condition, HERB's head moves at the same time and for the same distance as in the joint attention condition, but it moves laterally and remains oriented downward throughout this movement, so the gaze



(a) Neutral gaze. (b) Joint attention. (c) Mirrored gaze.

Figure 4: HERB began every trial in a neutral gaze position, looking at the block in its hand. Figures (b) and (c) show HERB in the two gaze conditions from the experiment.

appears non-social. Therefore, we control for total amount of head movement while manipulating whether the gaze is social or non-social. This manipulation explores whether social gaze before the suggestion affects how people respond to a counterintuitive suggestion.

3.1 Robot platform

Our robot HERB (Home Exploring Robot Butler) is a bimanual robot developed for assistive tasks in home environments [26]. HERB has two 7-DOF WAM arms, each with a 4-DOF BH8-series Barrett hand with three fingers. In this experiment, only the right arm was used. HERB's hand has a 6-axis force/torque sensor that can detect external forces applied to the joints, for instance when a participant gently pulls on an object in HERB's hand. Motion trajectories for picking up and handing over blocks were pre-planned using CHOMP [29] and played back during the experiment.

HERB also has a pan-tilt head outfitted with a Microsoft Kinect and a camera, though no real-time vision was used in this experiment. The front of the Kinect has two visible round cameras which serve as HERB's "eyes."

3.2 Procedure

Participants were randomly assigned to either the delay or no delay condition, and to either the joint attention or mirrored condition. There were 32 participants (18 females), eight in each of the four conditions, with a mean age of 34. Participants were recruited from the Pittsburgh area using an online participant pool website through Carnegie Mellon University. They were compensated $10 for their time.

Participants were told that they would play a sorting game with HERB. They were instructed to take the block from HERB's hand once HERB had extended the block to them. Participants were also told that HERB's head would move and that HERB may provide suggestions about how to sort the blocks, but that the final sorting method was up to them. Participants were not informed about the kinds of blocks they would be seeing, and the blocks were kept hidden under the table until HERB handed them to the participant.

On each trial, HERB picked up a block from below the table and handed it to participants by extending its arm forward while grasping the block. Following previous research [6], HERB's arm became fully extended to clearly communicate the handover.

HERB's hand contains a force sensor that identified when participants grasped the block during the handover. When the force sensor in the hand registered the participant's grasp on the block, HERB initiated a suggestion behavior by turning its head to one of the two boxes on the table. Depending on the delay condition, HERB either simultaneously released the block (no delay) or waited until its head was fully turned and then released the block (delay, Figure 2). Once the block was released, HERB withdrew its hand to begin the next trial.

Each participant engaged in five block handovers. The first two handovers involved solid color blocks, one of each color, both to familiarize participants with the task and to establish their preferred sorting method. The third handover involved the ambiguous block, presented with the first block's color on top. HERB always suggested that this block be sorted according to the color on the bottom, which violated the top-color sorting strategy, but was only a mildly counterintuitive suggestion because it did not violate the dominant color strategy (since there was no dominant color on the ambiguous block). The fourth block was again a solid colored block of the same color as the first block, intended to separate the test trials and to balance the number of blocks in each box as well as possible. For example, if the first block was yellow, the second block was blue, and the third block was ambiguous (presented with yellow on top), then the fourth block would be yellow again; the idea was that participants would sort the first and fourth blocks into the yellow box and the second and third blocks into the blue box, though this was not always the case. The final block was the semi-ambiguous block; this was always presented with the dominant color upward, to increase the saliency of the dominant color, but HERB always suggested sorting by the minority color. For example, if the fifth block was 70% blue, HERB oriented its head toward the yellow box.

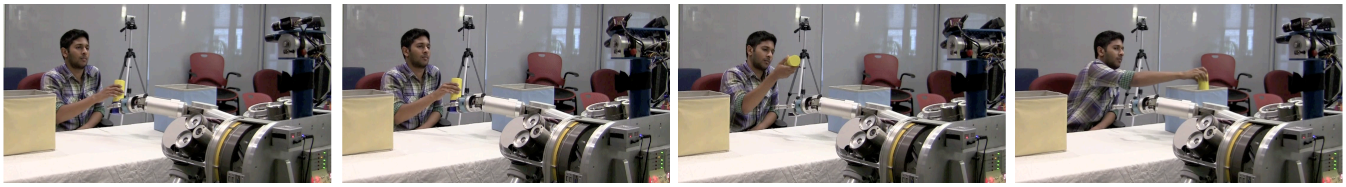
By handing the ambiguous and semi-ambiguous blocks over with HERB's suggested color on bottom, we made it as easy as possible for participants to use a sorting strategy that would conflict with HERB's suggestions in these cases. Therefore, we expect to see low compliance with HERB's suggestions in the absence of a manipulation.

First block color, dominant color of the semi-ambiguous block, and the arrangement of the boxes on the table were counterbalanced between participants. The experiment ran fully autonomously using pre-scripted trajectories for block pick-ups, block handovers, and head movements. There was real-time force feedback to measure when participants grasped the block during the transfer phase of the handover. The human-robot interactions lasted approximately 2 minutes and 20 seconds, though the particular amount of time varied by how long the participant took to sort the block.

3.3 Data collection

There are four data sources in this experiment. First, task performance was evaluated by whether participants followed HERB's suggestion on ambiguous and semi-ambiguous block trials. This provides a quantitative evaluation of compliance with HERB's counterintuitive suggestions.

Immediately after the interaction, participants completed written questionnaires that asked about their experiences and decisions during the task. These included free-response questions about whether they noticed suggestions from HERB and about their sorting strategy. Questionnaires contained specific questions about the ambiguous and semi-ambiguous blocks (represented with drawings), as well as Likert scale questions about HERB—rating features such as intelligence and friendliness—and about the collaboration—rating statements such as "I felt like HERB and I acted as a team."



(a) Delay

(b) No delay

Figure 3: A comparison of HERB's head and hand movements during the handover's transfer phase in delay and no delay conditions. In the no delay condition, HERB releases the block as it turns its head to a box. In the delay condition, the release occurs only after the head turn (frame 3).

After completing the survey, participants also engaged in a semi-structured interview with an experimenter. They were asked to explain their sorting of the ambiguous and semi-ambiguous blocks. They were also asked whether they noticed any suggestion behavior from HERB. The responses from these interviews are used to support participants' written responses in the surveys.

Finally, the interaction with HERB was video recorded, and these videos were coded for information such as amount of time spent looking at HERB's face and whether or not the participant looked at HERB's head as it executed a suggestion behavior. The videos were annotated by an independent coder naïve to the research hypothesis. We randomly selected 10% of the videos for validatation with a second coder; inter-coder agreement was 87% or higher. Because all of the coding measures were objective and inter-coder agreement was above the accepted 80% threshold, we feel confident analyzing the single coder's annotations.

4. RESULTS

This experiment yielded quantitative results from the task (such as the rate of participants complying with HERB's suggestion), self-reports in the form of Likert scales and free-responses on the post-task questionnaire and semi-structured interview, and objective observations of the interaction from the recorded videos (such as the amount of time participants spent looking at HERB's head).

Manipulation Check. To verify that HERB's sorting suggestion for semi-ambiguous blocks was counterintuitive, we analyzed the rate at which people chose the "counterintuitive" box when they were unaware of HERB's suggestion. Recognizing HERB's head movements as sorting suggestions significantly correlates with sorting the semi-ambiguous block as suggested (Pearson's $\chi^2(1, N = 32) = 11.567, p = 0.007$). Only one participant out of 14 sorted the semi-ambiguous block as HERB suggested without recognizing HERB's suggestion, verifying participants' bias against HERB's sorting suggestion and supporting the semi-ambiguous block as a valid manipulation to test for compliance.

Compliance: Semi-Ambiguous Block. The central research question is whether users comply with HERB's suggestion in the semi-ambiguous case. To test the effects of delay and gaze on correctness, we ran a factorial nominal logistic regression, which found that delay has a significant effect on compliance ($\chi^2(1, N = 32) = 6.77, p = 0.0092$). Without delay, only 19% of users sorted the semi-ambiguous block according to HERB's suggestion; delaying the release leads to 63% of users matching HERB's suggestion (Figure 5a).

When we analyze only participants who reported recognizing HERB's head movements as suggestions, the rate of compliance increases to 83% for participants in the delay condition and 33% for participants in the no delay condition (Figure 5b), with delay playing a significant role in this outcome: a nominal logistic regression for compliance with gaze and delay as factors, on only users who recognized HERB's head motions as suggestions, reveals a significant effect of delay ($\chi^2(1, N = 18) = 4.46, p = 0.0346$).

Compliance: Ambiguous Block. The ambiguous block represents a relatively low-conflict suggestion. Even though HERB always suggests sorting by bottom color, which violates the top-color strategy, most participants (59%) followed HERB's suggestion for sorting the ambiguous block.

An effect likelihood ratio test reveals a borderline significant effect for delay ($\chi^2(1, N = 32) = 3.632, p = 0.0567$), with 75% of participants in the delay condition following the ambiguous block suggestion, but only 56% of participants in the no delay condition following the suggestion (Figure 5c). There was no effect of gaze or an interaction.

Sorting ambiguous and semi-ambiguous blocks by HERB's suggestions are highly correlated (Pearson's $\chi^2(1, N = 32) = 9.85, p = 0.0017$). Ninety-two percent of users who sorted the semi-ambiguous block according to HERB's suggestion also previously sorted the ambiguous block according to HERB's suggestion.

Gaze. Gaze type (joint attention versus mirrored) did not significantly affect compliance, and there was no significant interaction effect. Similarly, the analysis of compliance in only participants who reported recognizing HERB's suggestions found no significant effect of gaze. Joint attention does correspond to a higher probability of following HERB's suggestion (44% with joint attention, 38% with mirrored), but the difference is not statistically significant.

Self-Reports. Participants' free responses on the questionnaire and interview revealed that 75% of participants in the delay condition and 38% of participants in the no delay

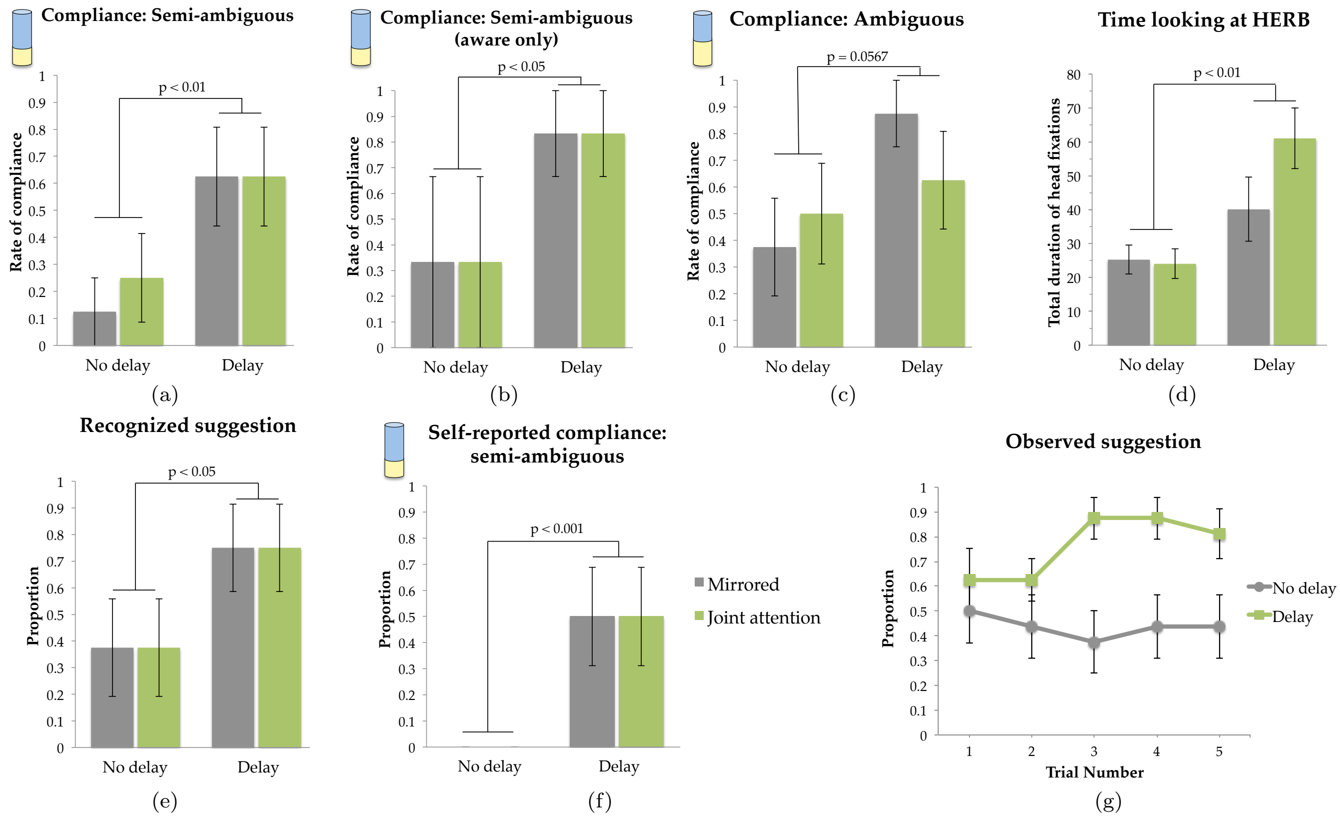


Figure 5: Results of the experiment from task responses, self reports, and video observations. Error bars indicate ± 2 SE.

condition noticed and interpreted HERB's head movements as sorting suggestions (Figure 5e). A nominal logistic regression with gaze and delay as factors shows that delay significantly affects whether participants thought HERB had a sorting suggestion ($\chi^2(1, N = 32) = 4.69, p = 0.0302$, Wald test $\chi^2(1, 32) = 4.32, p = 0.0377$). No significant effect was found for gaze, and no interaction was found.

None of the participants in the no delay condition thought they complied with HERB's suggestion on the semi-ambiguous block trial; they either did not state that HERB gave them a suggestion, or they explicitly stated that they did not follow HERB's suggestion (Figure 5f). In the delay condition, 50% of users explicitly stated that they chose their sorting strategy based on HERB's suggestion for the semi-ambiguous block. The effect is significant according to the effect likelihood ratio test ($\chi^2(1, N = 32) = 13.8, p = 0.0002$) but not according to the Wald test.

Video Coding. Videos were coded for how long participants looked at HERB's head, which reveal how much visual attention participants devoted to HERB's head during the task. Videos were also coded for events in which participants looked at HERB's head while HERB looked at one of the boxes, which indicate whether attention was directed to HERB's head at the right time to notice HERB's gaze cues.

A two-factor analysis of variance investigating the effect of delay and gaze on the total amount of time the participant looked at HERB's head showed a significant effect of delay ($F(1, 31) = 12.9828, p = 0.0012$). The handover delay more than doubled the mean looking time, from 24.7 seconds to 50.6 seconds (Figure 5d). There was no significant effect of gaze or an interaction.

To understand whether this additional time spent looking at HERB's head was useful, we ran a nominal logistic regression analyzing the effect of delay and gaze on whether the participant noticed HERB's suggestion in each trial, as measured by whether the participant looked at HERB's head while it was oriented toward one of the boxes (Figure 5g). The test found a significant effect of delay on noticing HERB's suggestion in the third and fourth trials, and a borderline significant effect in the fifth trial, but no significant effect of delay in the first or second trials ($\chi^2(1, N = 32) = 9.113, p = 0.003$ for trial 3, $\chi^2(1, N = 32) = 6.974, p = 0.008$ for trial 4, and $\chi^2(1, N = 32) = 4.993, p = 0.0254$ for trial 5, lowering the α to 0.01 for this analysis based on a Bonferroni correction, the most conservative control for multiple comparisons).

Given that HERB first presents the delay in the third trial, these results show that a deliberate delay led people to attend more to HERB's suggestions, even in subsequent trials when no delay was present (the fourth trial). The analysis did not find any effect of gaze on any trials.

5. DISCUSSION

Our results yield two main findings about the effects of deliberate handover delays.

RESULT 1. *Handover delays increased the amount of attention participants paid to HERB's head, which increased participants' awareness of HERB's non-verbal gaze cues.*

As predicted by hypothesis H1, deliberate delays increased the amount of time participants spent looking at HERB's head in general. This increase was not spurious: deliberate delays also increased time spent looking at HERB's head specifically when HERB made a suggestion. Therefore, a handover delay drew attention to HERB's head even though the head was not involved in the handover. These results are supported by both self-reports and video observations.

RESULT 2. *In addition to increasing recognition of HERB's suggestions, the handover delay also increased compliance with those suggestions.*

In our analysis of just the participants who reported recognizing HERB's head movements as suggestions, there was still a significant effect of delay on compliance. In other words, even once participants are aware that HERB is making a suggestion, they are still more likely to comply with that counterintuitive suggestion if the handover has a delay.

There are several interpretations for this second finding. The non-agentic explanation is that the delay drew peoples' attention to HERB's head which, because it was moving, increased the saliency of the suggestion behavior to the point where people followed it. This explanation does not require participants to attribute any kind of meaning to the handover delay. It is not wholly satisfactory, however, because when we exclude people who do not explicitly report seeing HERB's suggestion, there is *still* an effect of delay on compliance. Because these participants already notice HERB's head movements and explicitly interpret them as suggestions, the saliency of HERB's head movement seems unlikely to have a further effect.

A more agentic explanation is that HERB's handover delay was interpreted as *purposeful*, and that people were more likely to comply with HERB's suggestion when they believed that it came from an intentional agent. In order to test this explanation, we would need to measure peoples' attributions of agency to HERB before and after experiencing the delayed trials, an interesting point for future work.

In the current experiment, when the robot expressed a delay, there was always another meaningful channel of communication (eye gaze) to draw information from. However, it would be interesting to explore the effect of a delay when there is no other salient feature on which to focus attention. Perhaps in that case the delay would be interpreted as less agentic; the robot might even be seen as broken.

Mechanics of Deliberate Delays. We used a handover delay of one second in this study because that was the amount of time it took for HERB to turn its head toward a box. By the strength of the results, this duration was effective for drawing attention back to HERB's head.

A minority of participants in the delay case did not shift their attention in response to the delay; instead, these people seemed to focus more intently on pulling the block from HERB's hand. In these few cases, perhaps the delay was so unexpected that it served to draw attention to the hand, rather than release attention from it. More work is necessary to find the "sweet spot" where the delay is long enough to notice but not so long as to be problematic. This spot may also vary among people and depend on factors such as comfort with the robot.

Tasks for Investigating Compliance. The analysis supports our use of a semi-ambiguous block to investigate compliance. We found a strong correlation between noticing HERB's suggestion and sorting the block according to that suggestion. Of the 14 participants who did not report noticing HERB's suggestions, only one of them sorted the semi-ambiguous block in the same box that HERB suggested, emphasizing the counterintuitive nature of that suggestion. More users overall matched HERB's suggestion in the ambiguous case (59%) than in the semi-ambiguous case (41%). This is expected, as HERB's suggestion for the semi-ambiguous block conflicted with both the dominant color strategy and the top color strategy, whereas the ambiguous suggestion only conflicts with the top color strategy.

A block-sorting task is a useful proxy for other collaborative manipulation tasks because it involves many of the same behaviors in a simplified format. Our task included handovers, object manipulation, classification decisions, joint attention, referential gaze, and mutual gaze. It subtly addressed the issue of compliance and touched upon animacy and intentionality of robot agents. The task was performed in a constrained environment with high repeatability and few distractions, but it remained easy to understand and natural to complete.

Gaze. Our gaze manipulations had little effect on attention and compliance in this handover task. Though people could understand HERB's gaze to the box as a suggestion, there was no difference between joint attention and mirrored gaze conditions in terms of how much attention was directed at HERB's head or the rate at which participants complied with HERB's counterintuitive suggestions. Thus, hypothesis H2 was not supported.

While studies have shown that robot gaze is a strong social cue, many of these studies used tasks in which gaze was a primary component, such as conversation. In these situations, a person attends to the robot's gaze as part of the task, and therefore gaze cues may be more salient or useful.

Furthermore, HERB's "face", as seen in Figure 1a, is relatively abstract: the entire head consists of a flat platform with a pair of cameras for perception and a microphone. It is possible that HERB's non-anthropomorphic head affected how well people actually perceived joint attention, and that the joint attention condition would have yielded different results on a robot with more defined eyes.

In the delay condition, joint attention increased the amount of time spent looking at HERB's head to 61 seconds from 41 seconds for mirrored gaze. Joint attention also doubled the probability of a participant complying with HERB's suggestion in the no delay condition (from 12.5% to 25%). However, neither of these effects reached significance. More research is needed to understand how social gaze affects people in tasks where gaze is not a central component.

Future Work. Speech can be more precise and noticeable than gaze in many situations. Because this study focused on non-verbal communication (handover fluency and eye gaze), adding speech to the system would have been a confound with our current manipulations (handover delay and social gaze). However, future work on a robust human-robot handover system should incorporate spoken cues.

The decision to present the gaze cue during the handover transfer phase was based on a pilot and previous experience with HERB. Future studies can explore the effects of presenting the gaze cue at different points in the interaction.

Implications. The results reported in this paper provide insight into the design of effective robot-to-human handovers. When information needs to be conveyed during handovers, we suggest that seamlessness should be secondary to communication. For instance, by manipulating the force profile of the handover so that the robot deliberately delays releasing an object, robot designers can draw attention to important features like eye gaze and other non-verbal communication. This idea is in line with previous research that has shown that deliberately manipulating other aspects of a handover, like the spatio-temporal motion, can help convey information about the task [6]. The current work is novel because it uses a feature of the handover (the force profile) to convey information toward an unrelated mode of communication (eye gaze) about a subsequent task (block sorting).

6. ACKNOWLEDGMENTS

This work is supported by an Intel PhD Fellowship, NSF-IIS-1139078, NSF-IIS-1117801, NSF-EEC-0540865, ONR-N00014-12-1-0822, ONR-YIP 2012, and Intel Embedded Computing ISTC.

7. REFERENCES

[1] *A Roadmap for U.S. Robotics: From Internet to Robotics.* Robotics Virtual Organization, 2013.

[2] J. Aleotti, V. Micelli, and S. Caselli. Comfortable robot to human object hand-over. In *21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2012)*, pages 771–776, 2012.

[3] S. Andrist, T. Pejsa, B. Mutlu, and M. Gleicher. Designing effective gaze mechanisms for virtual agents. In *2012 ACM Annual Conference on Human Factors in Computing Systems (CHI '12)*, pages 705–714, 2012.

[4] A. Bicchi and V. Kumar. Robotic grasping and contact: A review. In *IEEE International Conference on Robotics and Automation (ICRA '00)*, volume 1, pages 348–353, 2000.

[5] J.-D. Boucher, U. Pattacini, A. Lelong, G. Bailly, F. Elisei, S. Fagel, P. F. Dominey, and J. Ventre-Dominey. I Reach Faster When I See You Look: Gaze Effects in Human-Human and Human-Robot Face-to-Face Cooperation. *Frontiers in Neurorobotics*, 6:1–11, 2012.

[6] M. Cakmak, S. S. Srinivasa, J. Forlizzi, and S. Kiesler. Human preferences for robot-human hand-over configurations. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '11)*, pages 1986–1993, 2011.

[7] M. Cakmak, S. S. Srinivasa, M. K. Lee, S. Kiesler, and J. Forlizzi. Using spatial and temporal contrast for fluent robot-human hand-overs. In *6th International Conference on Human-Robot Interaction (HRI '11)*, page 489, 2011.

[8] W. P. Chan, C. A. Parker, H. M. V. der Loos, and E. A. Croft. Grip Forces and Load Forces in Handovers: Implications for Designing Human-Robot Handover Controllers. In *8th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2013)*, pages 9–16, 2012.

[9] F. Dehais, E. A. Sisbot, R. Alami, and M. Causse. Physiological and subjective evaluation of a human-robot object hand-over task. *Applied Ergonomics*, 42:785–791, 2011.

[10] M. Diftler, J. Mehling, M. Abdallah, N. Radford, L. Bridgewater, A. Sanders, R. Askew, D. Linn, J. Yamokoski, F. Permenter, B. Hargrave, R. Platt, R. Savely, and R. Ambrose. Robonaut 2-The First Humanoid Robot in Space. In *2011 IEEE International Conference on Robotics and Automation (ICRA '11)*, pages 2178–2183, 2011.

[11] A. D. Dragan, K. C. T. Lee, and S. S. Srinivasa. Legibility and predictability of robot motion. In *8th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2013)*, pages 301–308, 2013.

[12] A. Edsinger and C. C. Kemp. Human-Robot Interaction for Cooperative Manipulation: Handing Objects to One Another. In *16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2007)*, pages 1167–1172, 2007.

[13] S. Glasauer, M. Huber, P. Basili, A. Knoll, and T. Brandt. Interacting in time and space: Investigating human-human and human-robot joint action. In *19th International Symposium in Robot and Human Interactive Communication*, pages 252–257, 2010.

[14] E. C. Grigore, K. Eder, A. G. Pipe, C. Melhuish, and U. Leonards. Joint Action Understanding improves Robot-to-Human Object Handover. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2013)*, 2013.

[15] M. Hayhoe and D. Ballard. Eye movements in natural behavior. *Trends in Cognitive Sciences*, 9(4):188–194, 2005.

[16] C.-M. Huang and A. L. Thomaz. Effects of responding to, initiating and ensuring joint attention in human-robot interaction. In *20th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2011)*, pages 65–71, 2011.

[17] M. Huber, M. Rickert, A. Knoll, T. Brandt, and S. Glasauer. Human-Robot Interaction in Handing-Over Tasks. In *17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*, pages 107–112, 2008.

[18] R. S. Johansson, G. Westling, A. Bäckström, and J. R. Flanagan. Eye-hand coordination in object manipulation. *The Journal of Neuroscience*, 21(17):6917–6932, 2001.

[19] M. K. Lee, J. Forlizzi, S. Kiesler, M. Cakmak, and S. Srinivasa. Predictability or Adaptivity? Designing Robot Handoffs Modeled from Trained Dogs and People. In *6th International Conference on Human-Robot Interaction (HRI '11)*, pages 179–180, 2011.

[20] B. Mutlu, J. Forlizzi, and J. Hodgins. A Storytelling Robot: Modeling and Evaluation of Human-like Gaze Behavior. In *6th IEEE/RAS International Conference on Humanoid Robots (Humanoids 06)*, pages 518–523. IEEE, 2006.

[21] B. Mutlu, T. Shiwa, T. Kanda, H. Ishiguro, and N. Hagita. Footing In Human-Robot Conversations : How Robots Might Shape Participant Roles Using Gaze Cues. In *4th ACM/IEEE International Conference on Human Robot Interaction (HRI 09)*, pages 61–68, 2009.

[22] B. Mutlu, F. Yamaoka, T. Kanda, H. Ishiguro, and N. Hagita. Nonverbal Leakage in Robots: Communication of Intentions through Seemingly Unintentional Behavior. In *4th ACM/IEEE International Conference on Human Robot Interaction (HRI 09)*, pages 69–76, 2009.

[23] H. Nguyen, C. Anderson, A. Trevor, A. Jain, Z. Xu, and C. C. Kemp. El-E: An assistive robot that fetches objects from flat surfaces. In *Robotic Helpers Workshop at HRI '08*, 2008.

[24] A. C. Pierno, M. Mari, S. Glover, I. Georgiou, and U. Castiello. Failure to read motor intentions from gaze in children with autism. *Neuropsychologia*, 44(8):1483–1488, 2006.

[25] E. A. Sisbot and R. Alami. A Human-Aware Manipulation Planner. *IEEE Transactions on Robotics*, 28(5):1045–1057, 2012.

[26] S. S. Srinivasa, D. Berenson, M. Cakmak, A. Collet, M. R. Dogar, A. D. Dragan, R. A. Knepper, T. Niemueller, K. Strabala, M. V. Weghe, and J. Ziegler. HERB 2.0 : Lessons Learned From Developing a Mobile Manipulator for the Home. *Proceedings of the IEEE*, 100(8):2410–2428, 2012.

[27] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, and S. S. Srinivasa. Learning the communication of intent prior to physical collaboration. In *21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2012)*, pages 968–973, 2012.

[28] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, S. S. Srinivasa, M. Cakmak, and V. Micelli. Toward Seamless Human-Robot Handovers. *Journal of Human-Robot Interaction*, 2(1):112–132, 2013.

[29] M. Zucker, N. Ratliff, A. Dragan, M. Pivtoraiko, M. Klingensmith, C. Dellin, J. A. D. Bagnell, and S. Srinivasa. Chomp: Covariant hamiltonian optimization for motion planning. *International Journal of Robotics Research*, 2013.