Methods for Expressing Robot Intent for Human–Robot Collaboration in Shared Workspaces

GREGORY LEMASURIER, GAL BEJERANO, VICTORIA ALBANESE, JENNA PARRILLO, and HOLLY A. YANCO, University of Massachusetts Lowell, USA
NICHOLAS AMERSON and REBECCA HETRICK, United States Air Force Academy, USA
ELIZABETH PHILLIPS, George Mason University, USA

Human–robot collaboration is becoming increasingly common in factories around the world; accordingly, we need to improve the interaction experiences between humans and robots working in these spaces. In this article, we report on a user study that investigated methods for providing information to a person about a robot's intent to move when working together in a shared workspace through signals provided by the robot. In this case, the workspace was the surface of a tabletop. Our study tested the effectiveness of three motion-based and three light-based intent signals as well as the overall level of comfort participants felt while working with the robot to sort colored blocks on the tabletop. Although not significant, our findings suggest that the light signal located closest to the workspace—an LED bracelet located closest to the robot's end effector—was the most noticeable and least confusing to participants. These findings can be leveraged to support human–robot collaborations in shared spaces.

CCS Concepts: • **Human-centered computing → HCI design and evaluation methods**; • **Computer systems organization** → *Robotic autonomy;*

Additional Key Words and Phrases: Signaling robot intent, feedback methods, collaborative robots

ACM Reference format:
Gregory LeMasurier, Gal Bejerano, Victoria Albanese, Jenna Parrillo, Holly A. Yanco, Nicholas Amerson, Rebecca Hetrick, and Elizabeth Phillips. 2021. Methods for Expressing Robot Intent for Human–Robot Collaboration in Shared Workspaces. *Trans. Hum.-Robot Interact.* 10, 4, Article 40 (September 2021), 27 pages.
https://doi.org/10.1145/3472223

1 INTRODUCTION

Workers in factory settings often have to collaborate with robots in close proximity. In a scenario where a human and robot are acting in the same space at the same time, it is important to avoid potentially dangerous collisions between the two workers. One way to avoid collisions is for the robot to signal its intent so that a human collaborator can plan their movements to avoid a collision

This research has been supported in part by the National Science Foundation under awards IIS–1426968 and IIS–1763469 and also supported in part by the Air Force Office of Scientific Research Grant 16RT0881f.
Authors' addresses: G. LeMasurier, G. Bejerano, V. Albanese, J. Parrillo, and H. A. Yanco, University of Massachusetts Lowell, One University Avenue, Lowell, Massachusetts, USA, 01854; emails: {Gregory_LeMasurier, Gal_Bejerano, Victoria_Albanese, Jenna_Parrillo}@student.uml.edu, holly@cs.uml.edu; N. Amerson and R. Hetrick, United States Air Force Academy, 2304 Cadet Drive, USAF Academy, Colorado, USA, 80840; emails: nick.amerson15@gmail.com, rmh0111@yahoo.com; E. Phillips, George Mason University, 4400 University Dr., Fairfax, Virginia, USA, 22030; email: ephill3@gmu.edu.
ACM acknowledges that this contribution was authored or co-authored by an employee, contractor, or affiliate of the United States government. As such, the United States government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for government purposes only.
© 2021 Association for Computing Machinery.
2573-9522/2021/09-ART40 $15.00
https://doi.org/10.1145/3472223

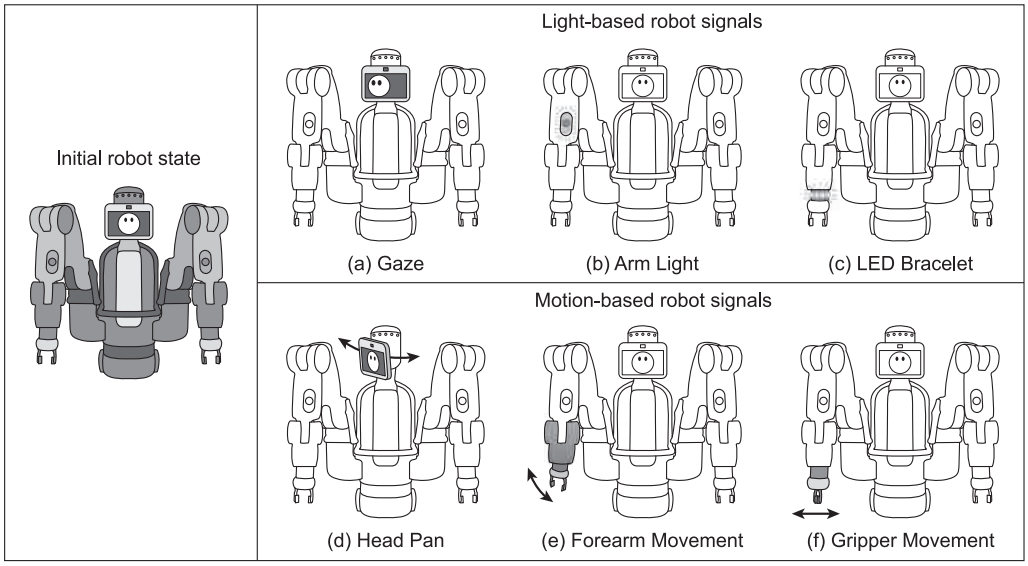

Fig. 1. Initial robot state (left) and the set of possible signals (right). Line drawings are used instead of photos to highlight the different signals. Each image shows the signal on the robot's right side; the portions of the robot being activated are shaded. A video depicting these signals can be found at https://youtu.be/YmpGw2kcEq0.

entirely. When necessary, the robot must be able to clearly show its intent in a way that the workers will notice and understand. Doing so can produce a safer and more efficient work environment for all.

To explore what methods could be most useful for conveying robot movement intent to a human collaborator, we developed multiple signals that allow the robot to express its intent to a person working in the same space, shown in Figure 1. We used a Baxter humanoid robot (human-likeness score = 27.30 [51]), created by Rethink Robotics for industrial pick and place tasks, as the robot platform for this experiment.

We designed six intent-signaling conditions to allow the robot to convey to a person which of its arms was going to move in a shared workspace while completing a collaborative task with a human partner. The six signals were designed to represent two super-ordinate categories, light-based signals and motion-based signals, each with three signals that varied in distance from the workspace. For each group, the farthest signals were located on the robot's head, the mid-distance signals were located on the robot's lower arm, and the closest signals were located on or near the robot's end effector. All signals were designed to require minimal to no modification of the robot. We did not design audio signals, because factories can have significant ambient noise, thus making audio signals impractical for our target applications. However, motion-based signals inherently had the secondary characteristic of making motor noise associated with the robot's movement.

Finally, we also included a no-signal (control) condition in which the robot moved its arms, without providing a signal to the human prior to doing so, while completing the task. In this condition, the robot paused briefly before executing its intended motion so that the robot's task completion times would be comparable across conditions.

Our results suggest that the light signal closest to the shared workspace (an LED Bracelet near the robot's end effector) was the most noticeable and least confused with other signals by participant. Although we observed that other motion signals were rated highly by participants, we concluded that the LED bracelet signal could provide additional benefits, because its signaling

modality (light) was distinct from the robot's normal motion, which may have been a confounding feature of some of the other signaling methods in our study. By identifying the signals that humans best understand and the optimal distance from the workspace for these signals, we hope to improve safety and efficiency for future human–robot collaboration, especially when collaborating in shared workspaces. The results of this work offer the following important contributions:

- In particular, we did not find in the literature instances in which the categories of motion and light-based intent signals, provided by robots, have been experimentally compared to one another. Thus, we designed, evaluated, and compared to one another a variety of motion and light-based intent signaling methods to be used when robots are collaborating in close proximity to humans.
- We discuss difficulties associated with creating signals that are interpretable and dissociated from a robot's task(s) in these environments.
- And, from our findings, we provide an empirically-based, practical recommendation for how an industrial-type robot can signal its intent to move to human collaborators in an easy and intuitive manner.

2 RELATED WORK

2.1 Human and Robot Coworkers in Collaborative Workspaces

As collaborative robots become more common, so do positions that require workers to interact with the robots. How do we ensure safe conditions for the humans while also optimizing the efficiency of both the human and robot workers? A variety of collaboration paradigms exist that all aim to optimize the hybrid factory. What differs between these paradigms is the degree to which the workspace and the work time is shared between humans and machines and the degree to which safety and efficiency goals are satisfied [27, 43].

Michalos et al. [43] explored several methods for improving workspace safety and efficiency. One method involved dividing the workspace into safety zones, used to regulate the behavior of the robotic workers based on the behaviors of the human workers; if a human was too close to the robot, then the robot would stop completely to avoid collisions. Another method granted a human collaborator some control over a robot, such as providing a compliant mode where the operator directs the robot's actions and the robot's force limit is reduced [43]. More recent work has utilized people predicting robot actions based on mental models that form after observing a robot. This behavior is harnessed by having a robot teach humans about its priorities so that a user mental model can form faster instead of through a long familiarization and inference process by the human [24].

The use of audio or tactile commands to help guide compliant robots has also been explored [67]. Audio commands, while intuitive for human operators, lack practicality in a factory environment filled with high-volume ambient noise. While collaborative situations such as these involve both shared workspace and shared work time, the relationship between the human and the robot is more similar to a supervisor–worker relationship, where the human has more power over the robot to guide it in its task. Our research is focused on improving a coworking relationship among human and robot collaborators. We find this important, because prior research by Hinds et al. [22] has explored manipulations of status where the robot was a subordinate, peer, or supervisor and found that people relied most on robot coworkers.

2.2 Signals of Human Intent

A human's expression of intent can be used by robots to predict and avoid areas that a person is likely to occupy in a coworking space. Humans' signals of intent can also be used as a basis for ways that robots can signal their own intent. For instance, models have been created that accurately predict a human's intention during a collaborative task by analyzing their gaze [23].

The use of intent signals, in particular gaze and gestures, have been analyzed in a user study where two people completed a handover task [60]. Algorithms have also been developed to improve a robot's awareness of and reaction to human intent [35, 69]. Adaptive algorithms have been shown to reduce the number of collisions and increase the overall task fluency between a human and robot that are working in a shared workspace [50]. Other groups have also investigated various human-aware navigation planners [28].

2.3 Signals of Robot Intent

Likewise, a robot can express its intent to convey its plan to collaborators. Prior research has investigated the balance between the robot conveying its goal ("legible") and the robot acting according to expectations ("predictable") [14]. Cha et al. [9] conducted a survey that focused on nonverbal signaling methods. Methods for robot intent can be classified into three categories: visual, motion, and audio signaling, where the majority of prior research on robot intent signaling has focused on visual signaling methods.

One form of visual intent signaling is through light signals. Light signals have been found to effectively express a robot's intent [58]. ModLight, a platform for light intent signaling has been developed to express a robot's movement and information about its current state through using a range of colors, speed, and brightness [11]. Other groups have also investigated using light signals to communicate a robot's state to a user [4]. Ambient light based signals have been found to be less distracting and less demanding than Graphical User Interfaces when communicating navigation information in autonomous vehicles [41]. One common form of light signaling is through blinker light signals, which are similar to the turn signals used in cars [16, 42, 57, 63]. Blinker light signals have been found to reduce the number of collisions between service robots and participants when lane changing in a hallway [16].

Robots have also been designed to communicate their intent through screen displays [57, 70]. Another popular visual signal is conveyed through the robot's gaze. One method of conveying gaze is through digital gaze [17, 39, 45, 54, 63], where light-based hardware such as arrays of lights or screen displays make the robot appear to be looking in a direction. Gaze has been found to improve task performance in a game where participants had to guess which object the robot was going to pick [48]. Other light patterns have also been designed to communicate a mobile robot's planned path [39, 63].

Another form of visual signaling includes projecting light into the environment to communicate information to surrounding workers. Some groups have investigated projecting a ray of light on the ground to express a robot's intended path to nearby people [39, 40, 57]. Other work has investigated using projectors to communicate the path that a mobile robot intends to take [12, 38, 39, 56, 71]. Projection mapping has been used to communicate a robot's internal states, such as detected objects in the robot's environment [20] and the object that the robot plans to grasp [20, 68]. Projected visual cues have also been used to communicate information such as the trajectory of industrial robot arms [56], the robot's workspace [3], as well as goal and task related information [3]. Some limitations of projectors include the need to be projected on flat surfaces and that they can only display two-dimensional (2D) positioning. These limitations make it hard to convey the intent of a robot that moves in three dimensions, such as a robot arm. Another problem that arises when using projectors for signaling is occlusion, if the robot or an object in the workspace were to block the projector, then the user would be unable to see the robot's planned trajectory.

In addition to physically projecting signaling cues into the environment, a robot can project its signaling cues virtually, using **augmented reality (AR)**. Walker et al. [66] investigated using virtual signaling in AR for aerial robots through virtual navigation points, arrows, gaze, and utilities. Williams et al. [72] used augmented reality as an effective way to provide deictic gestures that draw

robots' interlocutors' attention to different parts of the shared environment. Virtual signaling has also been used to communicate the planned trajectory of a robot's arm to an operator in AR [53].

As an alternative, robots can express their intent through motion signals as well. One existing signaling method is anticipatory motion, i.e., motion that conveys what the robot will do, to express intent signals [26]. A cost function has also been developed by Stulp et al. [61] to penalize joint execution time, task errors, and jerkiness that enables a robot to adapt its behavior so that a person understands its intent. Knight et al. investigated several movement based cues to indicate that the robot wanted to pass by [30]. Robots can also express intent through gestures and expressions [6]. Venture and Kulič analyzed various whole body expressive movements and established guidelines for movement based signals across various robot platforms and movement types [65]. Motion signals have been designed to communicate the flight path of assistive free flyers [62]. Several groups have also investigated signaling a mobile robot's intent using physical gaze, through head or physical eye movements [1, 2, 6, 29, 31, 42, 44, 48, 54].

Finally, audio signals can be used to express robot intent as well. Some work has investigated using a combination of audio and visual signals to express a robot's intent [10, 57]. Robot's can also express their intent through speech (e.g., References [13, 25]). Speech has also been used to communicate if the robot was accurately understanding and executing commands given by humans [52]. In addition to speech, audio tones and sounds can also be used to convey a robot's intent. Ulfvengren analyzed use of natural warning sounds as audio cues to indicate warnings to users in aviation [64]. Other work has also investigated using sonification to convey a robot's intent [7, 49]. Groups have also investigated using artificial motor sounds to communicate the intent of autonomous vehicles to pedestrians [46]. However, audio signaling is not practical in factory environments due to high-volume ambient noise common in these environments.

2.3.1 Limitations of Existing Methods. Many existing visual methods for signaling robot intent have involved providing signals to users working *around robots* sharing *common* spaces, not working *with robots* in *collaborative* spaces. For instance, many light-based signals have been used to keep robot and human movements through space independent of one another to avoid collision (e.g., depicting path and navigation plans). Light-based signals that communicate lane changes, in a hallway for instance, also rely on well-established understanding of how blinking indicator lights are used to signal lane changes for vehicles. Additionally, it is difficult to communicate 3D information using some light based signalling methods like projection. Both projection and AR methods can also be difficult to see in poor or changing lighting conditions. And, virtual signalling through AR also has the added limitation of requiring operators to wear equipment, such as glasses or a headset, to view the robot's signals. Without this equipment, no signals may be discernible from a robot that could potentially put people at risk. Finally, audio-based signals are not suitable for noisy environments common to many factory settings in which collaborative human–robot interactions are envisioned. Thus, when signalling methods are needed to communicate robot intent for humans and robots working together in space, when prior experience is not available, or in dynamic lighting or noise conditions, motion-based signals may be more suitable than light or audio-based signaling methods.

Although fewer prior research studies have investigated the usefulness of motion-based intent signals as compared to light-based intent signals, motion-based signals may provide utilities that light-based signals cannot. First, because motion-based signals are not directly tied to lights, projection systems, or external equipment, they are less likely to suffer from visibility problems associated with environmental lighting conditions or the use of external equipment. Further, they may be more noticeable and understandable in many settings, because they can be conveyed through constant external motion cues (e.g., like anticipatory motion) that may convey more information

about forthcoming intended motion. To illustrate, a light-based signal may convey forthcoming motion through a static or blinking light. Such conveyance requires observers to not only notice the light, but also translate its meaning as a cue that maps to intended motion. However, a motion-based cue is mapped to movement and thus does not require such a translation between signals and actions. However, people may perceive motion-based signals, where the robot is moving in close proximity to them, as more likely to cause human–robot collisions, which could counter some of the benefits of motion-based signaling methods.

3 HYPOTHESES

Four hypotheses motivated the current research. Each were previously registered in Reference [5].

Hypothesis 1: Intent shown through robot motion-based signals will be more effective at signaling robot motion intent than light-based signals that do not involve robot motion, where effective means seen/noticed and understood.

Hypothesis 2: Because effective signals should aid teammates in noticing and understanding how the robot is intending to move, more effective signals will allow the task to be completed faster.

Hypothesis 3: Intent shown through light-based signals will be more comfortable for participants than robot motion-based signals.

Hypothesis 4: Signals (either light-based or motion-based) located closer to the workspace will be more noticeable than those that are farther away.

4 METHOD

4.1 Participants

Ninety-two participants were recruited from the university and the surrounding city through fliers, email lists, and university social media groups. Eight participants were excluded from the dataset due to errors: Three for not understanding the instructions associated with the task and questionnaires (e.g., picking up more than one block at a time while sorting, filling out the NASA-TLX incorrectly), three because the robot's hardware failed while completing the task (e.g., Baxter's gripper got stuck), and two because their questionnaire data did not save properly. After exclusions, our final dataset included 84 participants, 12 in each of 7 conditions.

Participants' ages ranged from 18 to 69 ($M = 23.76, SD = 9.33$). Fifty-eight participants self identified as male and 26 as female. Two participants reported that they had a color vision deficiency and 4 participants reported that they had a hearing deficiency. Seventy-two participants said they were right handed, nine left handed, and three ambidextrous.

Our participants were also asked to report on a number of demographic variables to further characterize our sample or that could be used as covariates in subsequent data analyses. Specifically, participants were asked to report their agreement with the statement, "I have experience with robots" on a 7-point Likert-type scale that ranged from 1 (Strongly Disagree) to 7 (Strongly Agree). Twenty-seven participants reported above neutral on this scale, with mean agreement $M = 3.38, SD = 1.89$. Participants also reported their agreement with the statements "I frequently work on puzzles" and "I have experience with assembly tasks (for example, car repair)." Thirty-six participants reported above the neutral point on the scale for puzzle experience, $M = 3.96, SD = 1.54$ and 46 for assembly task experience, $M = 4.24, SD = 1.76$. Participants received a $10 gift card as compensation for their participation. The entire study took approximately one hour to complete and was approved by the university institutional review board.

4.2 Subjective Measures

Biographical data questionnaire: Participants were asked to fill out a biographical data questionnaire that asked for information including age, gender, handedness, known color vision

or hearing deficiencies, and prior experience with robots, working on puzzles, and assembly tasks.

Signal Effectiveness: Signal noticeability and understandability items: Signal effectiveness was operationalized in hypothesis H1 as how noticeable and understandable signals were to participants. Signal effectiveness is a subset of signal legibility. Dragan et al. defined legibility as being readable, anticipatory, and understandable [14]. Legibility also enables observers to make goal inferences [14]. In our experimental task, the robot and participant did not have shared goals. Therefore, providing goal inferences was not as important as indicating that the robot planned to move its arm. In order for signals that are effective to also be legible, they would have to enable goal inferences to be made. Thus, we used measures of signal noticeability and signal understandability to quantify signal effectiveness rather than signal legibility.

We asked participants about signal noticeability in two ways. The first measure presented participants with an image of the Baxter robot, similar to the leftmost depiction of Baxter in Figure 1, with instructions that read: "Please circle on the photo below what part of the robot was signaling to you how it would move." This measure was used to identify that the participants noticed the origin of the robot's intent signals. For the second, we asked participants to respond to the item "I noticed that the robot was signaling to me that it was planning to move" using a 7-point Likert-type scale that ranged from 1 (Strongly disagree) to 7 (Strongly agree).

Similarly, participants were asked to report how well they understood intent signals from the robot by responding to the item, "At any given moment I understood what the robot was doing," using the same 7-point Likert-type scale.

Comfort: Participants responded to a statement, "I felt comfortable with the robot," using a 7-point Likert-type scale that ranged from 1 (Strongly disagree) to 7 (Strongly agree).

NASA-TLX: The NASA-TLX [21] is a subjective measure of workload commonly used for assessing a variety of human-machine interfaces. Participants provided a rating of their perceived workload during a task across six sub-scales: mental demand, physical demand, temporal demand, effort, frustration, and performance. The first five sub-scales are measured from 0 (Low) to 100 (High), and performance is measured from 0 (Perfect) to 100 (Failure). The weighted measure of paired comparisons among the sub-scales was not included, because workload scores obtained with and without the weighted sub-scales are often correlated above $r = 0.90$ [47]. Thus, overall workload scores were calculated as the average of the six sub-scales after reverse coding scores on the performance dimension.

Robot role: Participants were asked to choose how they perceived the robot as a collaborator. When asked, "What was the robot's role in relation to you?," participants responded by choosing "No relation," "Manager," "Managed by me," or "Coworker."

System usability scale: Participants also filled out the **System Usability Scale (SUS)**. The SUS [8] assesses overall system usability by asking participants to rate 10 statements on a 7-point Likert scales ranging from "strongly disagree" to "strongly agree." The statements cover different perceptions of the system, such as complexity, consistency, and cumbersomeness. Each item on the SUS is combined to provide an overall usability score that can range from 0 (poor usability) to 100 (good usability).

Free response: We asked participants to provide us with open ended responses to a number of prompts including providing descriptions of what they thought the robot's strategy was for completing the task, as well as their strategy for completing the task, whether they thought the robot did anything unrelated to the task, and any final comments about the collaboration, the robot, or the experiment.

Ease of working with the robot: Finally, participants were asked to respond to three items using 7-point Likert-type scales, 1 (Strongly disagree) to 7 (Strongly agree), asking about the ease

of working with the robot (i.e., "It was easy to work with the robot," "The robot and I worked well together," and "The robot kept getting in my way"). Responses to these items were reversed scored when needed and averaged to provide an overall ease of working with the robot score, Cronbach's $a = 0.791$.

4.3 Objective Measures

Comfort working with the robot: To objectively evaluate participant comfort while working with Baxter, we recorded video of each participant interacting with the robot. We defined that participants who were more comfortable working with the robot would spend more time working on the experimental tasks while the robot was simultaneously working on its tasks. To complete the task in the most efficient manner possible, participants would need to collaborate with the robot, working simultaneously and in close proximity in the shared workspace, for instance by reaching for blocks across the shared workspace at the same time that Baxter was reaching for blocks, or working in the middle of the shared workspace at the same time that Baxter was moving around the space.

Two independent raters coded the video data for the specific timing of events throughout the human–robot collaboration. These included the task start and end times for the robot and the human completing the tasks, as well as the amount of time that the participants were idle while it was possible for them to complete part of the task (voluntary wait time). Participants working comfortably and efficiently with the robot would theoretically have little overall wait time—not pausing for Baxter, because they were uncomfortable moving close to the robot in the shared workspace, for instance. In addition, it was possible to work so quickly that any given participant would need to involuntarily wait on Baxter before completing the subsequent part of their task. To account for this, we asked our coders to record this involuntary wait time (if present) and we subtracted it from each participant's collaboration time. Using this event timing data we were able to calculate the proportion of time that participants were working while Baxter was moving given that they were capable of working during this time. This proportion is given by the following formula, where total possible collaboration time adjusts for involuntary wait time:

$$\frac{\text{Total possible collaboration time - Total voluntary wait time}}{\text{Total possible collaboration time}}.$$

Because our objective measure of comfort was derived by marking the timing of events in the video data, the rater agreement for coding when a particular event occurred depended somewhat on each coder's chosen video frame marking that event. An event like sorting the last block could span over multiple frames, for instance. Thus, a categorical or absolute measure of agreement (i.e., either an exact subsecond match between raters or not) like Cohen's Kappa would be artificially deflated and thus not be a reasonable capture of rater agreement. Intra-Class correlation coefficient handles this problem by providing a measure of agreement for continuous-level data rather than discrete data, and by accounting for error in the ratings [19]. Thus, we asked both coders to provide timestamps for 1/3 of the participants' video data. Then, inter-rater reliability was assessed via the Intra-Class correlation coefficient (3,1). Agreement ranged from 0.75 to 0.98, i.e., very good to nearly perfect agreement [32], for each of the time stamps that were used to calculate the objective measure of participant comfort. Because the coders showed near perfect agreement in their coding, we split the remaining 2/3 of participants' video data between the coders to derive the remaining objective comfort scores.

Total task time: From the timestamps provided by our raters we were also able to calculate the total time it took participants to complete the block sorting and pattern building tasks. Total task completion times were calculated by finding the amount of time participants took to complete the

sorting task (in seconds) and the amount of time participants took to complete the pattern building task (in seconds) and summing the two together.

4.4 Study Design

The study used a between-subjects design where each participant was randomly assigned to complete the experimental task under one of the six signal conditions—each representing a visual signal or a motion signal—or the control condition (no signal provided). A between-subjects design was chosen so that, in the post run questionnaire, we could directly ask participants if they had noticed a signal. This question would have biased participants to look for a signal if we had used a within-subjects design instead.

The manipulated variables were the type of signal (described below) and the distance of the signal from the workspace (closest signal on the gripper; mid-distance signal on the arm, farthest signal on the robot's head).

4.4.1 Light-based Intent Signaling Conditions.

(1) **Gaze:** Inspired by prior research showing the effectiveness of intent signaling using gaze for a flying robot [63], the screen on the robot's head displayed a face with eyes that shifted from the center to the right or left, toward the arm that the robot intended to move (Figure 1(a)). Once the robot's intended arm moved, the robot's gaze would shift back to the center. Only the image on the screen was changed; the position of the screen remained the same throughout this signal. While not physical eyes, this signal depicts moving video eyes in the form of shifting images, a type of digital gaze. This signal was the farthest visual signal from the shared workspace.
(2) **Arm Light:** The lights on the robot's arm would light up to signal which arm the robot was about to move (Figure 1(b)). These lights were blue, as provided by the manufacturer, and would turn on for a brief period of time before turning back off. The arm light would turn off shortly after the start of the robot's movement. The Arm Light signal was the mid-distance visual signal.
(3) **LED Bracelet:** We outfitted the Baxter robot with LED strip bracelets near the end effectors of each of the robot's arms. The LED Bracelet on the arm the robot intended to use would light up blue, and then that robot would move the arm (Figure 1(c)). The LED Bracelet would turn off shortly after the start of the robot's movement. LED Bracelet was the closest visual cue to the shared workspace.

4.4.2 Motion-based Intent Signaling Conditions.

(1) **Head Pan:** The robot panned its head from the center toward the arm that it planned to move (Figure 1(d)), which could be considered a physical gaze, like a person's head turn. After the arm started to move, the robot would pan its head back to the center. This signal was the farthest motion signal from the shared workspace.
(2) **Forearm Movement:** The robot twitched the elbow of the arm that it planned to move (Figure 1(e)), modeled after the anticipatory motion found effective for expressing robotic intent in a prior study [26]. This was the mid-distance signal.
(3) **Gripper Movement:** The robot opened and closed the gripper on the end effector of the arm that it intended to move (Figure 1(f)) prior to moving that arm. The Gripper Movement signal was the signal closest to the shared workspace.

Videos depicting the different signalling conditions used in this study can be found at https://youtu.be/YmpGw2kcEq0. There was also a no-signal (control) condition in which the robot moved its arms while completing the task without providing a signal to the human prior to doing so. In

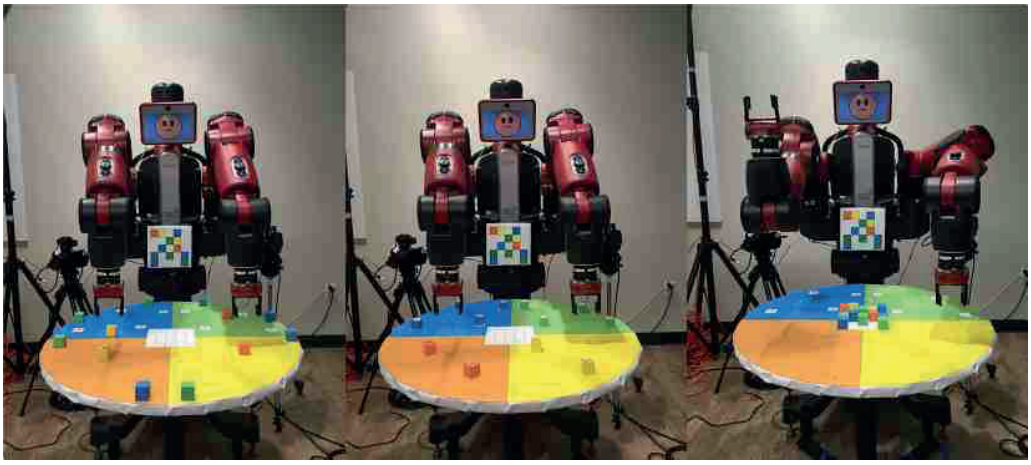

Fig. 2. Initial setup (left), all blocks sorted (middle), completed task with right hand up for a high-five (right).

this condition, the robot paused briefly before executing its intended motion so that the robot's task completion times would be comparable across conditions.

4.4.3 Experimental Task: Collaborative Block Sorting with the Baxter Robot. Participants were asked to collaborate with the Baxter robot to complete a block sorting and pattern building task on a shared table-top workspace (see Figures 2 and 3). Specifically, participants were tasked with sorting four green and four blue colored blocks by moving them into green and blue colored areas on the table. Each colored area had block-sized white squares labeled with a number between 1 and 4. Each colored block was labeled on the bottom with a number from 1 to 4. Participants were instructed that they were to sort the blue and green blocks one at a time, by both color and number by placing each numbered block in its corresponding color area and on its numbered square. Participants were also told that they could not begin sorting until Baxter had picked up its first block. This final instruction was created to standardize the start times across participants.

Simultaneously, Baxter was tasked with sorting two yellow and two orange blocks into corresponding colored areas on the table top, one at a time. The robot was given fewer blocks to sort, because it moves more slowly than a human. To sort the blocks, the robot moved its arms following a set of hard-coded waypoint positions to decrease variation in block sorting between subjects as a result of variation in the robot's autonomous motion planning. Once Baxter was done sorting, it would raise its end effector up into a position for a high-five, as shown in the picture on the right in Figure 2.

As soon as all the blocks of a single color (either Baxter's or the participant's) were sorted into their respective areas, the participant could begin working on the pattern portion of the task. The pattern portion of the task asked participants to move the colored blocks onto a grid in the center of the table top workspace and arrange the colored blocks to recreate an example abstract colored pattern. Participants were instructed to give the robot a high-five after completing the pattern or to tell the experimenter they were done.

4.5 Procedure

The experimenter conducting the study followed a script so that every participant received the same instructions. After entering the experimental room, participants were asked to read and sign an informed consent form and a photo/video release form. Then, participants were asked to fill out questionnaires on a computer to collect demographic data.

Fig. 3. A representative frame from the recorded quad camera view.

Participants read a sheet of instructions describing the block sorting task that they would complete with Baxter, then watched a four-minute video depicting the task rules by showing the task being played by a confederate participant and by a human acting out the robot's role. This video showed multiple acceptable strategies for completing the task (without emphasizing any) to avoid biasing the participants towards a particular strategy.

Next, participants engaged in a practice round of the sorting task with the experimenter, who played the role of the robot. This gave participants an opportunity to ask questions and for the experimenter to provide feedback on participant task completion to make sure that each participant understood the rules of the game. Participants then began the experimental phase of the study, where they collaborated with Baxter to sort blocks by color and number and then to recreate a pattern in the center of the shared workspace. After the experimental collaboration was complete, the participant filled out the subjective measures regarding their experience collaborating with the robot detailed above.

Finally, the experiment was recorded using four camcorders: one under the arm of the robot to capture the participants head-on, one to the right of the table (i.e., the robot's left) to capture the sides of the participant and the robot (i.e., the interaction), one to the back-left of the participants to capture the robot's actions, and a camera attached to a lighting rig above the table to capture a top down view of the shared workspace. Figure 3 shows the synchronized quadscreen view that was recorded featuring all camera views.

5 RESULTS

5.1 H1: Signal Effectiveness

H1 stated that motion-based signals would be more effective (i.e., noticeable and understood) than light-based signals. Three analyses of our measures used to operationalize signal effectiveness were conducted. First, a one-way between-subjects Kruskal–Wallis H-test with post-hoc analyses using the Tukey **Honest Significant Difference (HSD)** tests were conducted to look for statistically significant differences in participants' responses to the subjective signal noticeability item, "I noticed that the robot was signaling to me that it was planning to move" between the motion

signal, visual signal, and control conditions. Because our measure of signal noticeability consisted of a single item where participants used a Likert-type scale to respond, the non-parametric equivalent of the one-way ANOVA, the H-test, was used to test for significant differences between signaling conditions. To conduct this analysis, participant responses to the signal noticeability item in all of the motion-based conditions and in all of the light-based conditions were collapsed and then both were compared to participant responses in the control condition. Although the H-test was statistically significant, $H = 12.116, F(2, 81) = 6.767, p = 0.002, \eta^2 = 0.143$, it did not yield significant differences between the two super-ordinate types of conditions (i.e., motion vs. light). However, participants' mean scores on the signal noticeability item in both the motion-based ($M = 4.889, SD = 1.769$) and light-based ($M = 5.00, SD = 1.897$) signaling conditions were significantly higher than in the control condition ($M = 2.917, SD = 1.311$), $t = 3.338, p = 0.004$, Cohen's $d = 1.180$ and $t = 3.526, p = 0.003$, Cohen's $d = 1.174$, respectively. There were statistically significant differences in participant mean responses to the noticeability item between individual signaling conditions, $H = 20.670$, $F(6, 77) = 4.012, p = 0.002, \eta^2 = 0.238$. Post hoc comparisons (HSD) between all the signaling conditions revealed that there were three contrasts that were significantly different from one another. Participants in the LED Bracelet condition (light-based signal) agreed that they noticed the signal more than any other condition, $M = 5.833, SD = 1.583$ followed by the Head Pan condition $M = 5.583, SD = 1.676$ (motion-based signal), and the Forearm Movement condition (motion-based signal) $M = 5.083, SD = 1.564$. Mean scores in all three of these conditions were statistically significantly higher than participant responses in the control condition ($M = 2.917, SD = 1.311$), Cohen's $d's = 2.004$, 1.772, and 1.501, respectively, all $p's < 0.05$. The LED Bracelet, Head Pan, and Forearm Movement conditions were not significantly different from one another. Figures 4 and 5 depict mean scores as well as the proportion of participants' Likert scale responses on the noticeability item across signaling conditions.

Next, using participant responses to the signal noticeability image, we computed a confusion matrix of participant responses (Figure 6). This matrix describes the proportion of participants in each condition who circled the area on the robot that corresponded to their assigned condition as well as a range of areas that did not correspond to their assigned condition (e.g., participants in the Gripper Movement condition circling the wrist or lower elbow of the robot). More confusing conditions should reveal more evenly distributed proportions of participants circling across all the areas of the robots, whereas less confusing signal conditions should have high concentrations of participants circling the area of the robot that corresponds to their assigned condition and few participants circling anywhere else on the image of the robot.

The least confusing signal was the LED Bracelet signal given by the highest proportion of participants, 86.4%, correctly circling the robot's wrist on the signal noticeability image. In contrast, the Forearm Movement condition had the lowest proportion of participants (22.2%) who correctly selected the robot's lower elbow as the source of the robot's intent signal. For the Forearm Movement condition, it appeared that participants often confused the source of the signal as the robot gripper as 27% of participants in this condition incorrectly circled the gripper and not the robot's lower elbow. The ordering of the remaining conditions is given in Figure 6.

Finally, a one-way between-subjects H-test with HSD was conducted to test for significant differences in participant responses to the understandability item, "At any given moment, I understood what the robot was doing" between super-ordinate signaling conditions. Collapsing participant responses to the item in all of the motion-based signal conditions and in all of the light-based signal conditions and comparing them both to the control condition did not yield significant differences between the two super-ordinate types of conditions, $H = 0.944, F(2, 80) = 0.190, p = 0.827, \eta^2 = 0.005$. Further, the differences between all signaling conditions were not statistically significant, $H = 7.850, F(6, 76) = 1.304, p = 0.266, \eta^2 = 0.093$, although participants in the Gripper Movement

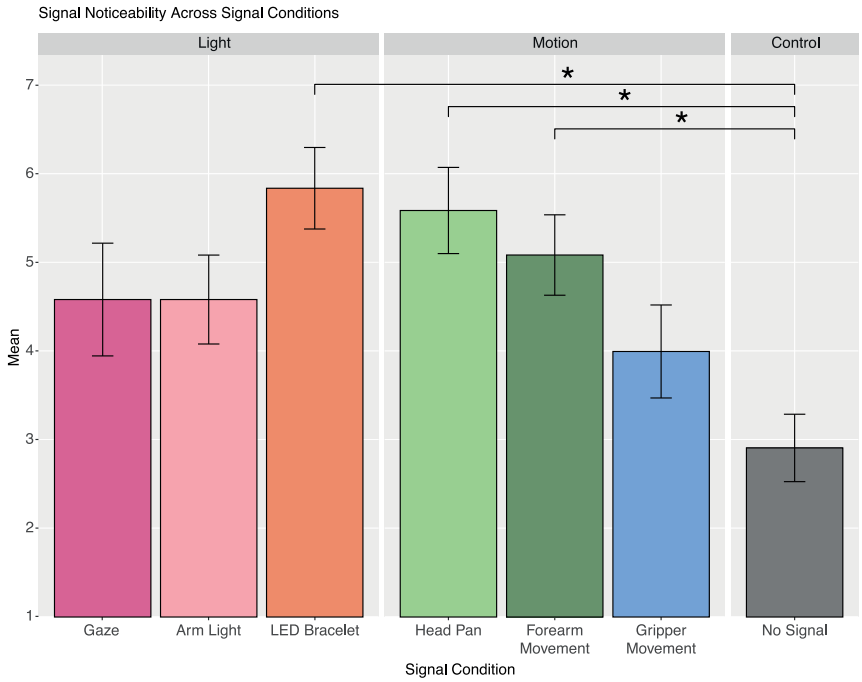

Fig. 4. Mean scores and standard errors across signaling conditions on the signal noticeability item, $^{*}p < 0.05$.

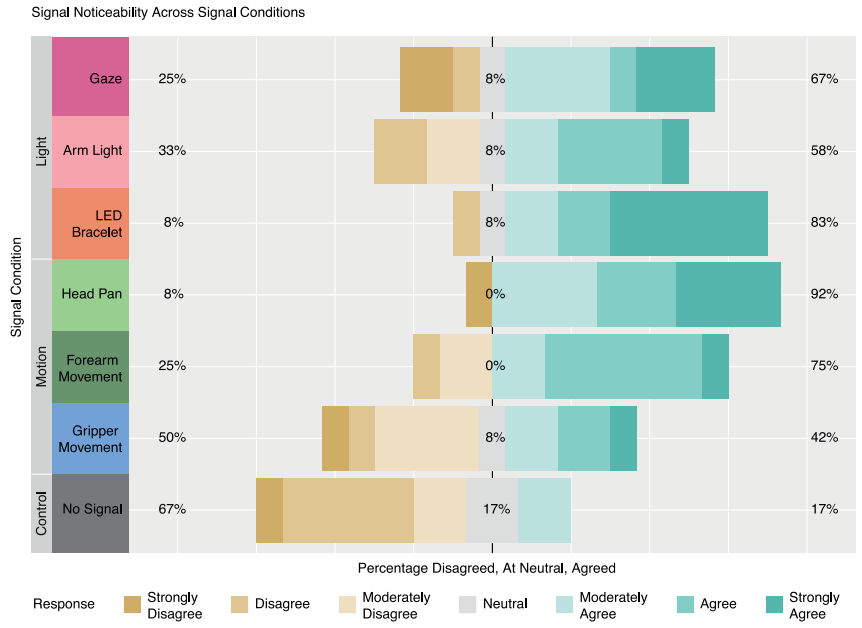

Fig. 5. Proportion of participant Likert scale responses to the noticeability item across signaling conditions. The percentages shown on the graph indicate the percent of participant responses that disagreed (left-hand side), were neutral (center axis), or agreed (right-hand side) in each of the signaling conditions.

	Circled Arm Light	Circled Wrists	Circled Head	Circled Lower Elbow	Circled Gripper	Didn't Notice	Circled Other
Arm Light	0.642	0.100	0.017	0.083	0.017	0.042	0.100
LED Bracelet	0.000	0.875	0.000	0.000	0.042	0.083	0.000
Gaze	0.017	0.017	0.625	0.100	0.142	0.083	0.017
Head Pan	0.033	0.089	0.417	0.172	0.172	0.083	0.033
Forearm Movement	0.017	0.100	0.083	0.267	0.267	0.167	0.100
Gripper Movement	0.058	0.149	0.000	0.218	0.426	0.083	0.065
No Signal	0.083	0.153	0.000	0.194	0.153	0.333	0.083

Fig. 6. Confusion Matrix showing the results of the article surveys on which participants circled the area(s) of the robot on which they perceived that an intent signal was given.

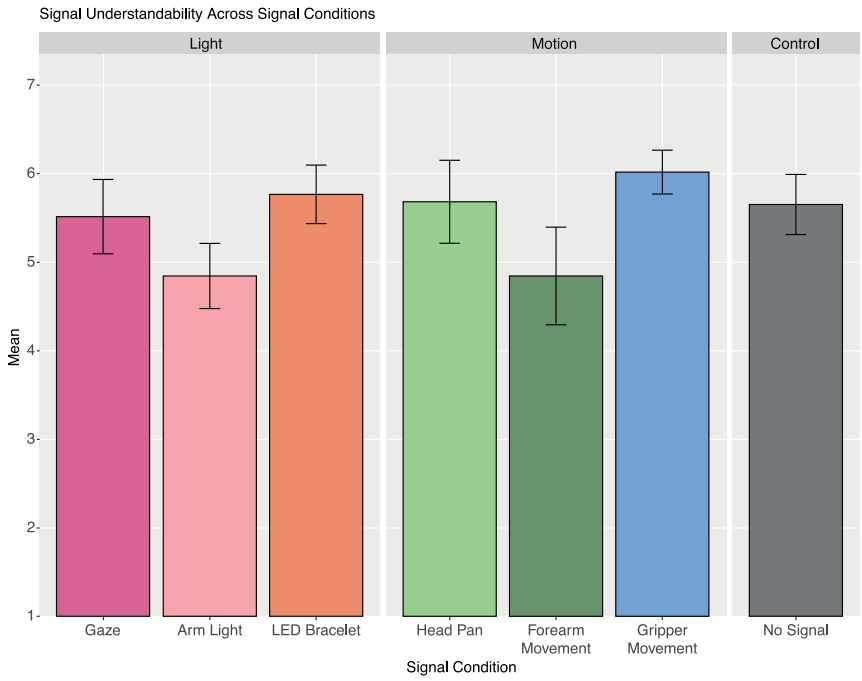

Fig. 7. Mean scores and standard errors across signaling conditions on the signal understandability item, $^{*}p < 0.05$.

condition reported the highest mean scores to the understandability item, ($M = 6.00, SD = 0.853$), followed by the LED Bracelet condition ($M = 5.750, SD = 1.138$), and the Head Pan condition ($M = 5.667, SD = 1.614$). See Figures 7 and 8.

These results support that the LED Bracelet signal (light-based) was the most noticeable and least confusing among our conditions, while the Gripper Movement signal (motion-based) was subjectively perceived to be the most understood by our participants. However, H1, which stated that motion-based signals would be more (a) seen and (b) understood than light-based signals, was not fully supported.

5.2 H2: Signal Effectiveness and Task Completion Time

Hypothesis H2 described that more effective signals would result in faster task completion times. To test H2, multiple linear regression was used with participant responses to the signal

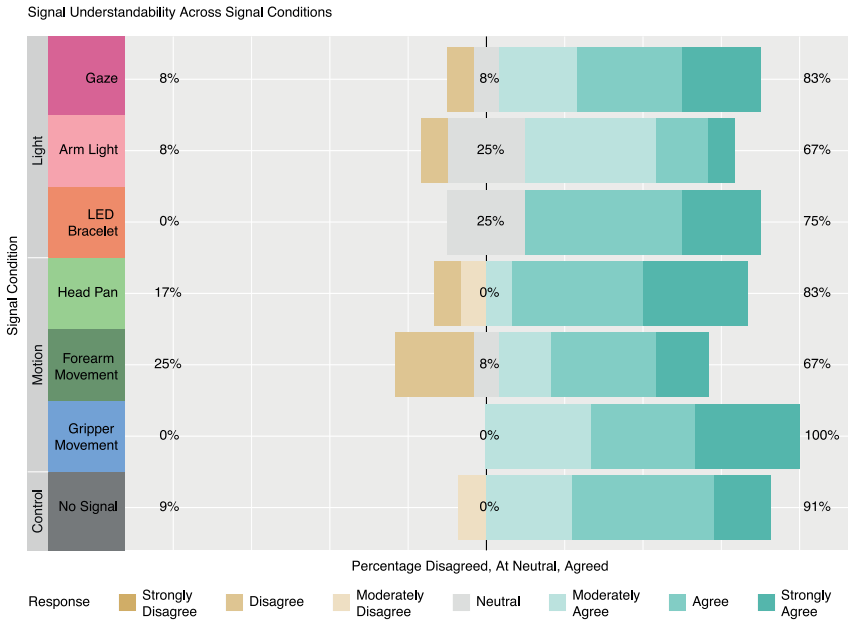

Fig. 8. Proportion of participant Likert scale responses to the understandability item across signaling conditions. The percentages shown on the graph indicate the percent of participant responses that disagreed (left-hand side), were neutral (center axis), or agreed (right-hand side) in each of the signaling conditions.

noticeability and understandability items (i.e., the two items used to measure signal effectiveness) entered as IVs and participant task completion times entered as the DV. The regression model as a whole was not significant. IV scores only explained 2% of the variance in task completion times, $R^2 = 0.002, F(2, 80) = 0.064, p = 0.938, r = 0.040$. Neither IV made a statistically significant unique contribution to the model; $b_{\text{noticeability}} = -0.40$; $t = 0.354, p = 0.725, r = -0.042$, $b_{\text{understandability}} = -0.012$; $t = 0.109, p = 0.913, r = -0.006$. As a result, H2 was not supported.

5.3 H3: Light-based Signals More Comfortable

Two groups of one-way tests with HSD were also conducted to look for significant differences in participant comfort between signaling conditions—one for the subjective measure of comfort and one for the objective measure.

To look for statistically significant differences between the super-ordinate signaling conditions (i.e. motion-based vs. light-based signals) on the subjective comfort measure, we again collapsed participants' responses to the comfort item across all the motion-based signal conditions and across all the light-based signal conditions and compared them both to the control condition using a Kruskal–Wallis H-test. The test revealed that there was not a statistically significant difference in scores on the subjective comfort item across super ordinate conditions, $H = 3.173, F(2, 81) = 1.684, p = 0.192, \eta^2 = 0.040$. Mean scores across the motion-based signal conditions were higher ($M = 5.861, SD = 1.125$) than in the light-based signal conditions ($M = 5.361, SD = 1.246$) and in the control condition ($M = 5.333, SD = 1.614$), although not significantly.

Without conditions collapsed, there was a statistically significant difference between signaling conditions on the subjective comfort measure $H = 15.072F = (6{,}77) = 2.704, p = 0.020, \eta^2 = 0.174$. Comparisons between all signals revealed that there were three contrasts that were significantly different from one another. Participants in the the Gripper Movement ($M = 6.083, SD = 0.793$)

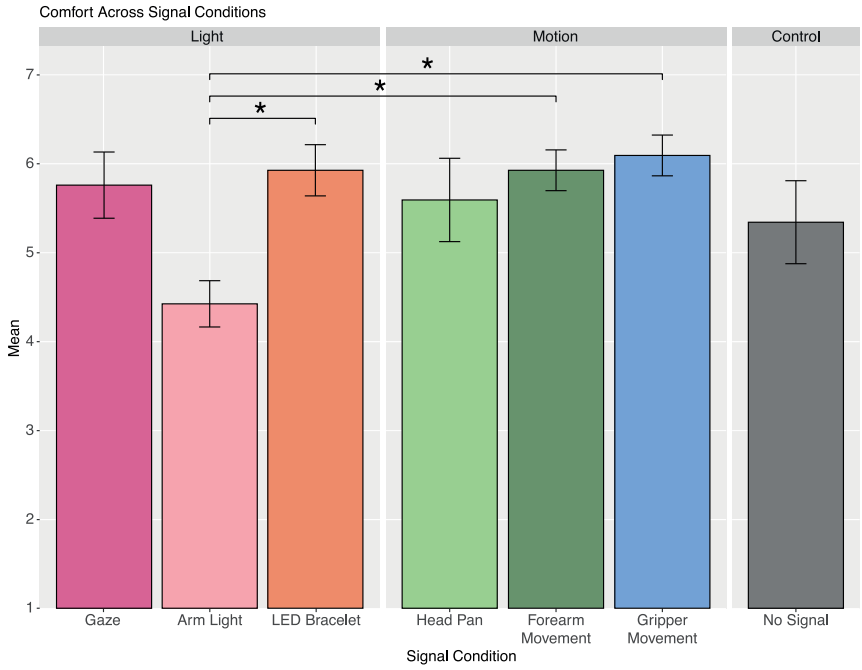

Fig. 9. Mean scores and standard errors across signaling conditions on the subjective comfort item, $^{*}p < 0.05$.

condition, followed by the Forearm Movement ($M = 5.917, SD = 0.793$) and LED Bracelet conditions ($M = 5.917, SD = 0.996$) reported that they felt significantly more comfortable than participants in the Arm Light condition ($M = 4.417, SD = 0.900$) Cohen's $d's = 1.965$, 1.768, and 1.580, respectively; all $p's < 0.05$. See Figures 9 and 10.

A one-way ANOVA was then used to test for significant differences between signaling conditions on the objective measure of participant comfort (i.e., the proportion of time participants spent working on experimental tasks while Baxter was simultaneously working on its tasks). The difference between super-ordinate conditions was not significant $F(2,81) = 1.684, p = 0.192, \eta^2 = 0.040$. The difference between individual signal conditions was also not significant, $F(6,77) = 0.375, p = 0.893, \eta^2 = 0.028$, and participants in the Gaze condition spent the largest proportion of time working while Baxter was working ($M = 0.762, SD = 0.269$), followed by the control condition ($M = 0.753, SD = 0.221$), and then the Forearm Movement condition ($M = 0.728, SD = 0.395$). H3, which proposed that light-based signals (on the whole) would be more comfortable than motion-based signals, was not supported.

5.4 H4: Closer Signals Are More Noticeable

A one-way Kruskal–Wallis H test with HSD was used to test hypothesis H4, which stated that participants would notice signals closer to the workspace more than signals further away from the workspace. Although the H-test was statistically significant, $H = 12.330, p = 0.006$, mean scores on the subjective noticeability item trended in the opposite of the hypothesized direction. Participants agreed that they noticed the robot's signals furthest from the workspace the most ($M = 5.083, SD = 1.976$), followed by the closest signals ($M = 4.917, SD = 1.909$), then the mid-distance signals ($M = 4.833, SD = 1.633$), and finally the no signal condition ($M = 2.917, SD = 1.311$). All distances of signals were significantly different from the no signal (control) condition

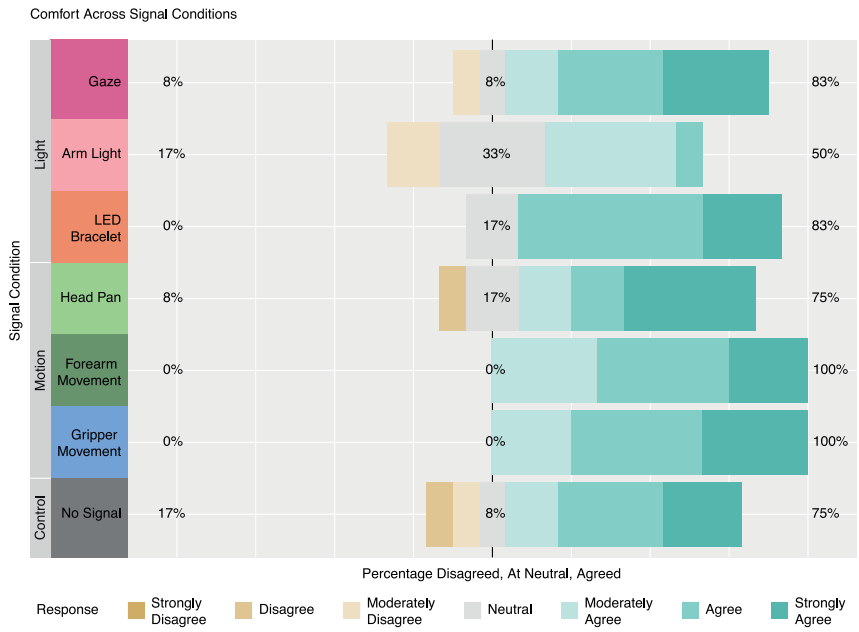

Fig. 10. Proportion of participant Likert scale responses to the subjective comfort item across signaling conditions. The percentages shown on the graph indicate the percent of participant responses that disagreed (left-hand side), were neutral (center axis), or agreed (right-hand side) in each of the signaling conditions.

($M = 2.917, SD = 1.311$), Cohen's $d's = 1.211$, 1.150, and 1.248, respectively, all $p's < 0.05$, but not from one another. See Figure 11. Because mean scores trended in the opposite of the hypothesized direction (i.e., signals furthest away from the workspace were subjectively reported to be most noticeable), H4 was not supported.

5.5 Supplementary Analyses

A series of supplementary analyses were conducted to look for significant differences between signaling conditions in participant perceptions of ease of working with the robot, perceptions of system usability, and cognitive workload as a result of working with the robot.

5.5.1 Ease of Working with the Robot. Two one-way ANOVAs with HSD were conducted to test for differences in participants' perceptions of the ease of working with the robot between signaling conditions. We found a significant main effect across signal conditions, $F(6,77) = 3.864, p = 0.002, \eta^2 = 0.231$. Post hoc comparisons revealed that participants in the Forearm Movement condition ($M = 6.000, SD = 0.682$) perceived it to be significantly easier to work with the robot than participants in the Arm Light ($M = 4.028, SD = 0.969$) and control conditions ($M = 4.556, SD = 1.166$), Cohen's $d's = 2.355$ and 1.512, $p < .001$ and $p = .027$ respectively. Participants in the Forearm Movement condition reported the highest mean scores on our measure of ease of working with the robot ($M = 6.000, SD = 0.682$), followed by the Gripper Movement condition ($M = 5.333, SD = 1.378$) and the LED Bracelet condition ($M = 5.139, SD = 1.000$), although the differences between these conditions were not statistically significant. See Figure 12.

Significant differences in scores on the measure of ease of working with robot were found between the super-ordinate conditions as well, $F(2, 81) = 6.674, p = 0.008, \eta^2 = 0.113, M_{visual} = 4.69$,

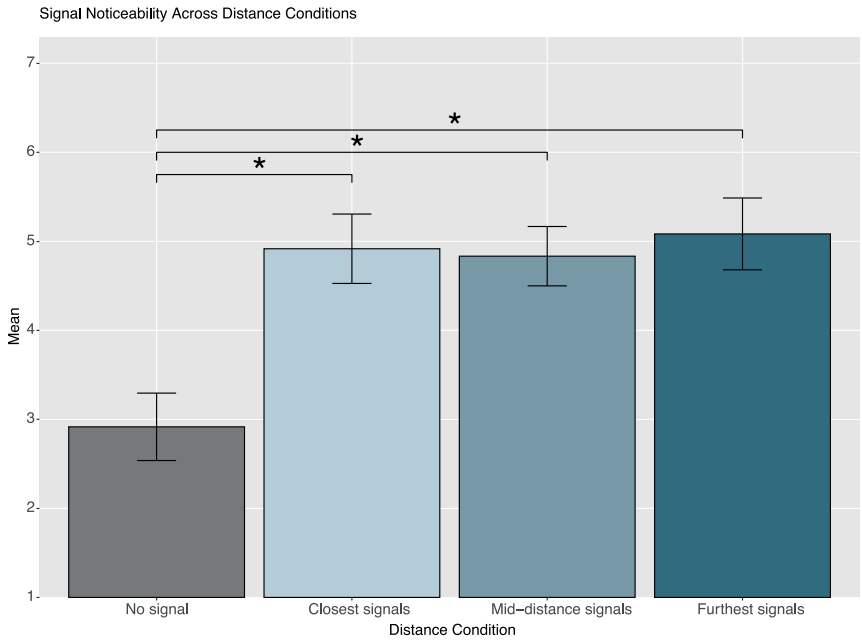

Fig. 11. Mean scores and standard errors across signal distance conditions on the signal noticeability item, $^{*}p < 0.05$.

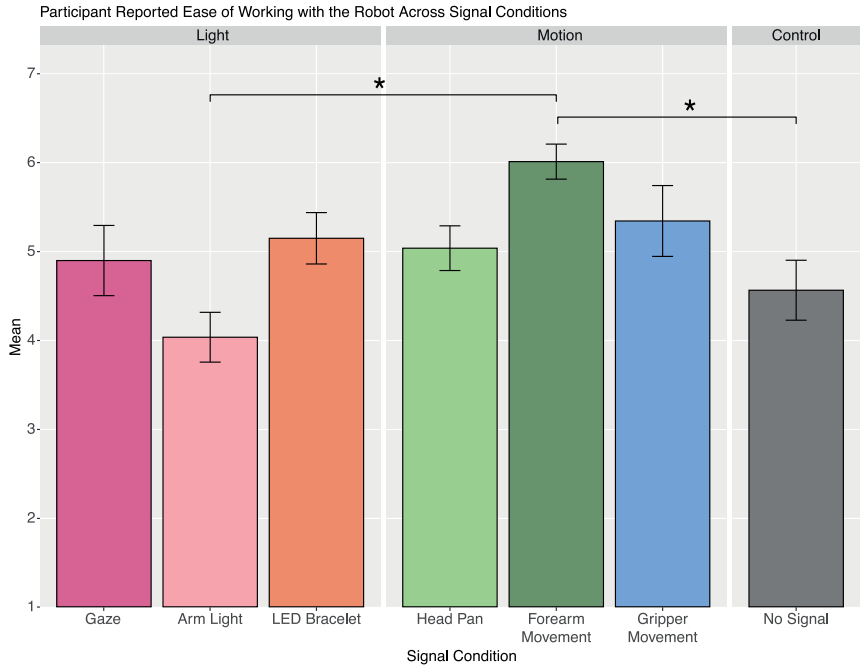

Fig. 12. Mean scores and standard errors across signaling conditions on the measure of ease of working with the robot, $^{*}p < 0.05$.

$SD = 1.20$, $M_{motion} = 5.45$, $SD = 1.07$, where motion-based signals were reported to be easier to work with than light-based signals.

5.5.2 System Usability. To test for significant differences in participant perceptions of system usability across the signaling conditions, two one-way ANOVAs with HSD were used; one for SUS scores across super-ordinate conditions, and one between all the signaling conditions. There were no significant differences between super-ordinate conditions $F(2, 81) = 2.054, p = 0.135, \eta^2 = 0.048$, or between individual signaling conditions, $F(6, 77) = 1.796, p = 0.111, \eta^2 = 0.123$. Further, scores trended such that motion-based signals were associated with the highest average usability scores ($M = 70.853, SD = 15.035$), followed by light-based signals ($M = 65.645, SD = 12.714$). Participants in the no signal condition reported the lowest perceived system usability scores ($M = 62.768, SD = 13.981$), with participants in the Forearm Movement condition reporting the highest usability scores ($M = 72.738, SD = 11.531$), followed by the Gripper Movement condition ($M = 72.589, SD = 16.857$) and then the LED Bracelet condition ($M = 72.440, SD = 10.818$).

5.5.3 Cognitive Workload. Two one-way between-subjects ANOVA with HSD were again used to test for significant differences in scores on the NASA-TLX measure of cognitive workload between the super-ordinate light-based, motion-based, and control conditions. There were no significant differences in cognitive workload between the super-ordinate conditions, $F(2, 81) = 2.002, p = 0.142, \eta^2 = 0.047$, or between individual signaling conditions, $F(6, 77) = 1.123, p = 0.357, \eta^2 = 0.080$. However, scores trended such that motion-based signals were associated with the lowest average workload scores ($M = 30.481, SD = 17.280$), followed by the light-based signals ($M = 36.750, SD = 14.952$). Participants in the no signal condition reported the highest cognitive workload scores ($M = 40.611, SD = 23.762$). Comparing individual signaling conditions, participants in the Forearm Movement condition reported the lowest cognitive workload scores ($M = 25.883, SD = 14.131$), followed by the Gripper Movement condition ($M = 30.944, SD = 14.840$), the Arm Light condition ($M = 33.944, SD = 17.286$), the Head Pan condition ($M = 34.667, SD = 22.074$), and the LED Bracelet condition ($M = 34.927, SD = 14.888$).

5.5.4 Short Answer Themes. The short answer responses on the survey were used to identify any themes in participant perceptions of collaborating with the robot that were not captured in our other dependent measures. The first such theme is a contradiction of beliefs, a cognitive dissonance, namely, that the robot avoided the participant (an incorrect belief), but that the participant actively stayed out of the way of the robot. Of the 84 participants, 19 indicated that they believed that the robot sensed and worked around them (a false statement). Of these 19 participants, 13 also indicated that they had to work around the robot. This means that more than two thirds of the participants who believed the robot avoided them had evidence of the contrary, which they ignored. One respondent wrote, "I would say the robot's strategy was looking for any empty open space and trying not to get in my way ([it] wasn't very good at it though)." It seems that some participants assumed that the robot (or the robot's designer, as some indicated) was at least trying to be courteous.

A second theme was that an understanding of the rules helped the participant predict the robot's actions. In the responses, 17 participants indicated that they used the collaboration task rules (e.g., Baxter sorting only the orange and yellow blocks, having to sort all of one color before starting on the pattern in the middle of the table) to assume where the robot was going. For example, one participant wrote, "I did [understand what the robot was doing] because I had read the purpose of the study and also because of the [game rules] video. If I came in cold to the experiment I would not have fully understood the intention of the robot." An additional 5 participants indicated that

predicting the robot became easier as the task went on, due to the reduced number of possible robot moves.

6 DISCUSSION

6.1 Signal Noticeability

H1, which stated that motion-based signals provided by the robot would be more effective, meaning both seen more and understood more by participants, was not supported. Results supported that the LED Bracelet signal (light-based) was the most noticeable and least confusing among our signal conditions (although it was not statistically signficant, it was more noticeable than the other signals). while the Head Pan signal was the most noticed motion-based signal and the Gripper Movement signal was the least noticeable but reportedly most understood overall. Some interesting insights for these results can be gained by looking at participant responses to the signal noticeability image of the Baxter robot and the resulting signal confusion matrix. Specifically, the LED Bracelet signal was reported to be noticed more by participants and was the most frequently correctly identified signal when we asked participants to indicate the source of signals on a physical representation of the robot. The LED Bracelet signal also represented one of the closest workspace signals that was hypothesized in H4 to be more noticeable than signals at other distances.

Although H4 was not supported, the wording of H4 necessitated that we collapse motion-based and light-based signals in each of the three distance conditions. Doing so combined participant responses to the noticeability item in the LED Bracelet condition with participant responses in the Gripper Movement condition, which had the lowest scores on the noticeability item aside from the control condition ($M = 4.00, SD = 1.81$) and which the confusion matrix revealed was often confused with the Forearm Movement signal. Isolating scores in the LED Bracelet signal condition from the Gripper Movement signal condition, provides support for this claim. Specifically, after doing so, mean scores on the noticeability item were highest in the LED bracelet signal condition ($M = 5.83, SD = 1.59$), and higher than the composite conditions of the mid and farthest distance signals as well as the Gripper Movement and Control conditions. This finding provides partial support for H4, which stated that the signals closer to the workspace (e.g., LED Bracelet) would be more noticeable than signals further away from the workspace.

We also realized that the noticeability of the Gripper Movement signal was likely confounded and confused with the robot's normal task motion and possibly with other signals. To illustrate, the opening and closing of the gripper was used to signal the robot's intention to move; however, the opening and closing of the gripper was also a necessary and frequent part of the robot picking up, moving, and sorting blocks during the collaboration task. Thus, participants may have not been able to distinguish between when the gripper was providing an intention signal or completing some other element of the sorting task. This confounding likely accounted for the Gripper Movement condition's relatively low scores on the signal noticeability measures, as well as both the Gripper Movement and the Forearm Movement conditions' relatively high scores on other outcomes of interest like understandability, comfort, usability, and cognitive workload. Participants may not have perceived the Gripper Movement intent signal as a signal at all, possibly, because it was so similar to other elements of the robot's task.

The Forearm Movement signal was also likely confused with the Gripper Movement signal. The Forearm Movement signal was least likely to be identified correctly in the confusion matrix for the signal noticeability image. Although the origin of the robot's signaling motion in the Forearm Movement condition was the robot's lower elbow, the resulting motion would manifest primarily in the movement of the robot's end effector—analogous to a human bending at the elbow and their hand following in motion as a result. Thus, it is understandable that participants in the Forearm

Movement and Gripper Movement conditions often confused the origin of the signals given in their respective conditions. For many of our subjective measures, participants scores in the Forearm Movement and Gripper Movement conditions were also very similar, for example SUS scores, ease of working with the robot scores, and the scores on the subjective comfort item. In participant free responses, participants also often used language that acknowledged a signal but could apply to either Forearm Movement or Gripper Movement, for example, "Before the robot moved his arm, he flexed his claw" was provided by a participant in the Gripper Movement condition.

Finally, if we continue with the inference that participants confused signals, then it would lend additional explanation for why participants reported that they noticed the signals furthest away from the workspace more than signal's closer to the workspace. Signals closest to the workspace (especially motion) were likely to be confused with the robot's normal motion in completing the sorting task, and mid-distance signals (like Forearm Movement) were likely confused with close signals like the Gripper Movement. In contrast, signals on the robot's head were completely disassociated from the physical task of sorting blocks and other signals. Thus, signals that were dissociated from the robot's task movement (e.g., signals on the robot's head or the LED Bracelet) may have been more salient when recognized and distinguishable from the rest of the robot's motion.

6.2 Signal Understandability

If it is true that the Gripper Movement signal was perceived to be part of the robot's normal motion, then it is not entirely surprising that participants would find the gripper opening and closing to be understandable as part of the robot's tasking, while not recognizing it as a signal. Although the Gripper Movement signal (motion-based) was perceived to be the most understandable, which would partially support H1, it is probably not appropriate to argue that the reason why the Gripper Movement was understandable was due to the participants truly perceiving the signal as discussed previously. It is likely the case that the Gripper Movement seemed understandable, because it was not perceived to be distinct from the robot's other movements. In contrast, the LED Bracelet signal was perceived to be relatively understandable—second in the ordering of conditions, while simultaneously being noticed the most and the least confused with other signals.

6.3 Task Performance

Regarding task completion times, neither measure of signal noticeability and understandablity considered together, nor treated independently were significantly predictive of participant task completion times.

6.4 Participant Comfort

H3 stated that light-based signals would be more comfortable for participants than motion-based signals. H3 was not supported as motion-based signals were subjectively reported to be more comfortable than light-based signals, although mean scores in these super-ordinate conditions were not statistically significantly different from one another. Further, the Gripper Movement and Forearm Movement conditions were reported by participants to be the most comfortable and there were no significant differences between conditions on the objective measure of comfort, potentially for the reasons already mentioned. Again, the Gripper Movement being reported as the most comfortable is also not surprising if we adhere to the explanation that participants perceived the gripper motion as consistent with the robot's normal motion and that the Forearm Movement and Gripper Movement signals were often confused with one another by participants. Thus, it is a relatively positive finding that the LED Bracelet signal was rated as just as comfortable (i.e., identical mean scores) as the Forearm Movement signal.

Regarding our objective measure of comfort in H3, it is very possible that our objective operationalization of comfort was not a good measure of participant comfort and potentially not a measure of comfort at all. It is possible that our measure was actually a reflection of confusion or (in)effectiveness of signals. Participants may have been spending more time working with the robot (i.e., objective comfort), because they did not understand the ways in which the robot would move—attending to and interpreting a signal may have slowed them down overall, while not paying attention to the robot allowed them to focus on their physical task of sorting their blocks, and thereby spend more time working while robot was working.

6.5 Summary and Implications of Robot Signal Findings

To ease comparison of conditions across all of the dependent measures, we provide two summaries of findings in Tables 1 and 2. Table 1 presents each condition's relative rank (e.g., 1st most, 2nd highest, etc.) on each dependent outcome, and Table 2 provides the sum of each rank to give an overall score across outcomes. Lower scores indicate collectively better outcomes across the measures. The LED Bracelet signal had relatively low scores (i.e., high ranks) on several of our outcome measures (e.g., noticeability, understandability comfort, usability, ease of working with the robot), and thus the lowest summary score of the conditions. Further, some participants in the LED Bracelet condition even commented on factors that impacted comfort, with one participant reporting, "The movements were predictable so I was comfortable working with it," and another saying, "The robot had a comforting design and looked peaceful. I did not fear for my safety as well."

By contrast, although the Gripper and Forearm Movement conditions ranked well on a number of our outcome measures, it is difficult to determine why, or to rule out the possibility that these two conditions were confused and confounded with each other and/or other motions of the robot. Often mean scores in the LED Bracelet condition were not statistically significantly different from the Forearm Movement and Gripper Movement conditions as well. These findings have potentially positive design implications. Namely that the LED Bracelet signal was collectively the best signal across conditions and has the practical benefits of being close to the shared workspace as well as distinct from robot motion, and correctly identifiable as a signal, which may make the signal salient to human collaborators if deployed in industrial settings. A challenge for deployment, then, will be to ensure that the signal is both salient and understandable by human collaborators.

7 ETIQUETTE AND COLLABORATION

Responses to the open-ended prompts provided cursory evidence that some participants believed that Baxter had knowledge of their movement and that the robot actively avoided collisions with them, which in reality, the robot was not doing. It seemed that some participants were inferring that the robot (or its designers) intended for the robot to act in accordance with politeness or etiquette norms (e.g., staying out of the way of the participant, turn taking) when collaborating on the task. These findings are consistent with a body of literature that suggests that people indeed treat and expect robots to be social interactors [18, 34], and may trigger social schemas for interaction [33, 59]. Further, researchers in the human–robot interaction community have highlighted the importance of and need for research focused on developing norm understanding in robot collaborators [36, 37, 55], which can make robots more predictable, and as a result trustworthy. However, false perceptions of robot norm following may also create problems associated with overtrusting the robot, believing it is capable of completing tasks it is unable to, and relying on the robot inappropriately. We note this here to keep in mind when designing robot signals in follow-up studies.

Table 1. A Comparison of the Six Intent Signals and the Control Signal across Dependent Measures

LED Bracelet	Arm Light	Gaze
1st most noticeable** (Section 5.1)	4th most noticeable (Section 5.1)	5th most noticeable (Section 5.1)
1st least confusing (Section 5.1)	2nd least confusing (Section 5.1)	3rd least confusing (Section 5.1)
2nd most understood (Section 5.1)	6th most understood (Section 5.1)	5th most understood (Section 5.1)
3rd highest subjective comfort †† (Section 5.3)	7th highest subjective comfort ‡ (Section 5.3)	4th highest subjective comfort (Section 5.3)
4th highest objective comfort (Section 5.3)	7th highest objective comfort (Section 5.3)	1st highest objective comfort (Section 5.3)
3rd highest ease (Section 5.5.1)	7th highest ease † (Section 5.5.1)	5th highest ease (Section 5.5.1)
3rd highest system usability (Section 5.5.2)	7th highest system usability (Section 5.5.2)	5th highest system usability (Section 5.5.2)
5th lowest cognitive workload (Section 5.5.3)	3rd lowest cognitive workload (Section 5.5.3)	7th lowest cognitive workload (Section 5.5.3)

Gripper Movement	Forearm Movement	Head Pan
6th most noticeable (Section 5.1)	3rd most noticeable** (Section 5.1)	2nd most noticeable** (Section 5.1)
4th least confusing (Section 5.1)	7th least confusing (Section 5.1)	5th least confusing (Section 5.1)
1st most understood (Section 5.1)	7th most understood (Section 5.1)	3rd most understood (Section 5.1)
1st highest subjective comfort †† (Section 5.3)	2nd highest subjective comfort †† (Section 5.3)	5th highest subjective comfort (Section 5.3)
5th highest objective comfort (Section 5.3)	3rd highest objective comfort (Section 5.3)	6th highest objective comfort (Section 5.3)
2nd highest ease (Section 5.5.1)	1st highest ease †† (Section 5.5.1)	4th highest ease (Section 5.5.1)
2nd highest system usability (Section 5.5.2)	1st highest system usability (Section 5.5.2)	4th highest system usability (Section 5.5.2)
2nd lowest cognitive workload (Section 5.5.3)	1st lowest cognitive workload (Section 5.5.3)	4th lowest cognitive workload (Section 5.5.3)

Control
7th most noticeable (Section 5.1)
6th least confusing (Section 5.1)
4th most understood (Section 5.1)
6th highest subjective comfort (Section 5.3)
2nd highest objective comfort (Section 5.3)
6th highest ease (Section 5.5.1)
6th highest system usability (Section 5.5.2)
6th lowest cognitive workload (Section 5.5.3)

** Statistically significant relative to control.
† Statistically significant relative to Forearm Movement, lower than control.
†† Statistically significant relative to Arm Light.
‡ Statistically significant relative to Gripper Movement, Forearm Movement, and LED Bracelet.

Table 2. A Summary of the Six Intent Signals and the Control Signal across Dependent Measures

	LED Bracelet	Gripper Movement	Forearm Movement	Head Pan	Gaze	Arm Light	Control
Most Noticeable (Section 5.1)	1**	6	3**	2**	5	4	7
Least Confusing (Section 5.1)	1	4	7	5	3	2	6
Most Understood (Section 5.1)	2	1	7	3	5	6	4
Highest Subj. Comfort (Section 5.3)	3††	1††	2††	5	4	7‡	6
Highest Obj. Comfort (Section 5.3)	4	5	3	6	1	7	2
Highest Ease of Use (Section 5.5.1)	3	2	1††	4	5	7†	6
Highest System Usability (Section 5.5.2)	3	2	1	4	5	7	6
Lowest Cognitive Workload (Section 5.5.3)	5	2	1	4	7	3	6
Total	**22**	**23**	**25**	**33**	**35**	**43**	**43**

Lower scores indicate more favorable performance.
**Statistically significant relative to control.
†Statistically significant relative to Forearm Movement.
††Statistically significant relative to Arm Light.
‡Statistically significant relative to Gripper Movement, Forearm Movement, and LED Bracelet.

8 LIMITATIONS

There are a number of limitations to this study that should be mentioned. Prior to participant recruitment, we conducted a power analysis that informed our goal to recruit 140 participants to complete the study across the seven independent conditions (6 signals and 1 control). Unfortunately, we were unable to meet this recruitment goal; thus, our study was under powered. A post hoc power analysis using G*Power [15] revealed that across our main analyses we achieved power between $\beta = 0.06$ and $\beta = 0.30$ with an average beta of $\beta = 0.17$. A larger sample would have allowed us to be more confident in the stability of the mean differences found between signaling conditions. Additionally, we recognized that the motion of the Forearm Movement and Gripper Movement conditions were confounded, and that the Gripper Movement may also have been too

similar to the robot's normal task motion, which likely added to participant confusion, and makes it difficult to make concrete design recommendations about these conditions even though mean scores on many of our outcomes of interest favored these conditions. Our objective measure of comfort was operationalized as the proportion of time participants were working in close proximity to Baxter while Baxter was moving given that participants were capable of doing so. This operationalization was likely not a valid capture of participant comfort in the way we intended. Our measure of task time did not seem to be sensitive to our experimental manipulations as well. It is also possible that participants may have been able to complete the sorting task(s), using different strategies that would not necessarily reflect participant comfort, signal effectiveness, or other features of our experimental manipulations. Participants could possibly have been able to complete their tasks very quickly for instance without ever attending to or understanding an intent signal from the robot. Finally, we used a variety of single-item subjective dependent measures for outcomes of interest. Creating scales consisting of multiple items could help us to assess how reliably participants were evaluating our experimental constructs and creating and validating such measures would be helpful for the HRI community on the whole.

9 CONCLUSIONS

The LED Bracelet added to Baxter was the most noticeable and least confusing signaling method for our participants. This was likely aided by the fact that the LED Bracelet was the closest in proximity to the shared workspace *and* was distinct from the robot's other movement. However, the Arm Light signal was rated relatively poorly on almost all of our outcome measures and in some instances, was rated lower than having no signal at all (e.g., comfort, understandability). This may be because it was difficult to see on the robot. The Forearm Movement, which like the Arm Light is a mid-distance signal, was likely confused with the Gripper Movement signal. Adding mid-distance signals to collaborative robots like Baxter may not be the most practical for facilitating comfort and collaboration. The Gripper Movement and Forearm Movement signals were rated relatively high on several measures; however, our study does not allow us to discern whether or not this is due to participants actually perceiving the signal or confusing the signal for something else. As such, the most practical recommendation based on these data would be that the addition of light signals near the workspace and origin of motion (like an LED Bracelet) may provide a benefit to humans collaborating in shared spaces with robots like industrial settings.

ACKNOWLEDGMENTS

Thanks to Adam Norton for his assistance with the figures, including the design of Figure 1. Thank you to the reviewers for their contributions to improving the quality of this article. The views expressed in this document are the authors and do not reflect the official position of the USAF Academy, USAF, or U.S. Government.

REFERENCES

[1] Henny Admoni, Anca Dragan, Siddhartha S. Srinivasa, and Brian Scassellati. 2014. Deliberate delays during robot-to-human handovers improve compliance with gaze communication. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*. 49–56.

[2] Pourya Aliasghari, Moojan Ghafurian, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. 2021. Effects of gaze and arm motion kinesics on a humanoid's perceived confidence, eagerness to learn, and attention to the task in a teaching scenario. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*. 197–206.

[3] Rasmus S. Andersen, Ole Madsen, Thomas B. Moeslund, and Heni Ben Amor. 2016. Projecting robot intentions into human environments. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'16)*. IEEE, 294–301.

[4] Kim Baraka, Ana Paiva, and Manuela Veloso. 2016. Expressive lights for revealing mobile service robot state. In *Proceedings of the 2nd Iberian Robotics Conference (Robot'15)*. Springer, 107–119.

[5] Gal Bejerano, Gregory LeMasurier, and Holly A. Yanco. 2018. Methods for providing indications of robot intent in collaborative human-robot tasks. In *Companion of the ACM/IEEE International Conference on Human-Robot Interaction*. 65–66.

[6] Cynthia Breazeal, Cory D. Kidd, Andrea Lockerd Thomaz, Guy Hoffman, and Matt Berlin. 2005. Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 708–713.

[7] Roberto Brensin, Emma Frid, Adrian B. Latupeirissa, and Claudio Panariello. 2021. Robust non-verbal expression in humanoid robots: New methods for augmenting expressive movements with sound. In *Proceedings of the Sound in Human-Robot Interaction Workshop*.

[8] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usabil. Eval. Industr.* 189, 194 (1996), 4–7.

[9] Elizabeth Cha, Yunkyung Kim, Terrence Fong, Maja J. Mataric, et al. 2018. A survey of nonverbal signaling methods for non-humanoid robots. *Found. Trends Robot.* 6, 4 (2018), 211–323.

[10] Elizabeth Cha and Maja Matarić. 2016. Using nonverbal signals to request help during human-robot collaboration. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'16)*. IEEE, 5070–5076.

[11] Elizabeth Cha, Tushar Trehon, Lancelot Wathieu, Christian Wagner, Anurag Shukla, and Maja J. Matarić. 2017. Modlight: designing a modular light signaling tool for human-robot interaction. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'17)*. IEEE, 1654–1661.

[12] Ravi Teja Chadalavada, Henrik Andreasson, Robert Krug, and Achim J. Lilienthal. 2015. That's on my mind! robot to human intention communication through on-board projection on shared floor space. In *Proceedings of the European Conference on Mobile Robots (ECMR'15)*. IEEE, 1–6.

[13] Joe Crumpton and Cindy L. Bethel. 2016. A survey of using vocal prosody to convey emotion in robot speech. *Int. J. Soc. Robot.* 8, 2 (2016), 271–285.

[14] Anca D. Dragan, Kenton C. T. Lee, and Siddhartha S. Srinivasa. 2013. Legibility and predictability of robot motion. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI'13)*. 301–308.

[15] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 2 (2007), 175–191.

[16] Rolando Fernandez, Nathan John, Sean Kirmani, Justin Hart, Jivko Sinapov, and Peter Stone. 2018. Passive demonstrations of light-based robot signals for improved human interpretability. In *Proceedings of the 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'18)*. IEEE, 234–239.

[17] Stephen M. Fiore, Travis J. Wiltshire, Emilio J. C. Lobato, Florian G. Jentsch, Wesley H. Huang, and Benjamin Axelrod. 2013. Toward understanding social cues and signals in human–robot interaction: Effects of robot gaze and proxemic behavior. *Front. Psychol.* 4 (2013), 859.

[18] Kerstin Fischer. 2011. Interpersonal variation in understanding robots as social actors. In *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI'11)*. IEEE, 53–60.

[19] Kevin A. Hallgren. 2012. Computing inter-rater reliability for observational data: An overview and tutorial. *Tutor. Quant. Methods Psychol.* 8, 1 (2012), 23.

[20] Zhao Han, Alexander Wilkinson, Jenna Parrillo, Jordan Allspaw, and Holly Yanco. 2020. Projection mapping implementation: Enabling direct externalization of perception results and action intent to improve robot explainability. In *the AAAI Fall Symposium on The Artificial Intelligence for Human-Robot Interaction (AI-HRI'20)*.

[21] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology*. Vol. 52. Elsevier, 139–183.

[22] Pamela J. Hinds, Teresa L. Roberts, and Hank Jones. 2004. Whose job is it anyway? a study of human-robot interaction in a collaborative task. In *Human-Computer Interaction*, Vol. 19. 151–181.

[23] Chien-Ming Huang, Sean Andrist, Allison Sauppè, and Bilge Mutlu. 2015. Using gaze patterns to predict task intent in collaboration. *Frontiers in Psychology* 6 (July 2015), 1049.

[24] Sandy H. Huang, David Held, Pieter Abbeel, and Anca D. Dragan. 2019. Enabling robots to communicate their objectives. In *Autonomous Robots*, Vol. 43. 309–326.

[25] H. Huttenrauch and Kerstin Severinson Eklundh. 2003. To help or not to help a service robot. In *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN'03)*. IEEE, 379–384.

[26] Michael J. Gielniak and Andrea L. Thomaz. 2011. Generating anticipation in robot motion. In *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'11)*. 449–454.

[27] A. Verl, J. Krüger, and T. K. Lien. 2009. Cooperation of human and machines in assembly lines. In *CIRP Annals*. 628–646.

[28] Harmish Khambhaita and Rachid Alami. 2017. Assessing the social criteria for human-robot collaborative navigation: A comparison of human-aware navigation planners. In *Proceedings of the 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'17)*. IEEE, 1140–1145.

[29] Harmish Khambhaita, Jorge Rios-Martinez, and Rachid Alami. 2016. Head-body motion coordination for human aware robot navigation. In *Proceedings of the 9th International Workshop on Human-Friendly Robotics (HFR'16)*.

[30] Heather Knight, Timothy Lee, Brittany Hallawell, and Wendy Ju. 2017. I get it already! the influence of chairbot motion gestures on bystander response. In *Proceedings of the 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'17)*. IEEE, 443–448.

[31] Raymond H. Cuijpers and Marco AMH Knops. 2015. Motions of robots matter! The social effects of idle and meaningful motions. In *International Conference on Social Robotics*. Springer, 174–183.

[32] Terry K. Koo and Mae Y. Li. 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 2 (2016), 155–163.

[33] Min Kyung Lee, Sara Kiesler, and Jodi Forlizzi. 2010. Receptionist or information kiosk: How do people talk with a robot? In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*. 31–40.

[34] Manja Lohse. 2012. Treating robots as social beings-A matter of personal preconceptions or interpersonal alignment?. In *Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'12)*. IEEE, 839–844.

[35] Jim Mainprice and Dmitry Berenson. 2013. Human-robot collaborative manipulation planning using early prediction of human motion. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'13)*. 299–306.

[36] Bertram F. Malle and Matthias Scheutz. 2014. Moral competence in social robots. In *Proceedings of the IEEE International Symposium on Ethics in Science, Technology and Engineering*. IEEE, 1–6.

[37] Bertram F. Malle, Matthias Scheutz, and Joseph L. Austerweil. 2017. Networks of social and moral norms in human and robot agents. In *A World with Robots*. Springer, 3–17.

[38] Takafumi Matsumaru. 2006. Mobile robot with preliminary-announcement and display function of forthcoming motion using projection equipment. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communicatio (RO-MAN'06)*. IEEE, 443–450.

[39] Takafumi Matsumaru, Kazuya Iwase, Kyouhei Akiyama, Takashi Kusada, and Tomotaka Ito. 2005. Mobile robot with eyeball expression as the preliminary-announcement and display of the robot's following motion. *Auton. Robots* 18, 2 (2005), 231–246.

[40] Takafumi Matsumaru, Takashi Kusada, and Kazuya Iwase. 2006. Mobile robot with preliminary-announcement function of forthcoming motion using light-ray. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 1516–1523.

[41] Andrii Matviienko, Andreas Löcken, Abdallah El Ali, Wilko Heuten, and Susanne Boll. 2016. NaviLight: investigating ambient light displays for turn-by-turn navigation in cars. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 283–294.

[42] Alyxander David May, Christian Dondrup, and Marc Hanheide. 2015. Show me your moves! conveying navigation intention of a mobile robot to humans. In *Proceedings of the European Conference on Mobile Robots (ECMR'15)*. IEEE, 1–6.

[43] George Michalos, Sotiris Makris, Panagiota Tsarouchi, Toni Guasch, Dimitris Kontovrakis, and George Chryssolouris. 2015. Design considerations for safe human-robot collaborative workplaces. In *Procedia CIRP 37*. 248–253.

[44] Masahiko Mikawa, Yuriko Yoshikawa, and Makoto Fujisawa. 2018. Expression of intention by rotational head movements for teleoperated mobile robot. In *Proceedings of the IEEE 15th International Workshop on Advanced Motion Control (AMC'18)*. IEEE, 249–254.

[45] Nicole Mirnig, Nicole Perterer, Gerald Stollnberger, and Manfred Tscheligi. 2017. Three strategies for autonomous car-to-pedestrian communication: A survival guide. In *Proceedings of the Companion of the ACM/IEEE International Conference on Human-Robot Interaction*. 209–210.

[46] Dylan Moore, Rebecca Currano, and David Sirkin. 2020. Sound decisions: How synthetic motor sounds improve autonomous vehicle-pedestrian interactions. In *Proceedings of the 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 94–103.

[47] William F. Moroney, David W. Biers, F. Thomas Eggemeier, and Jennifer A. Mitchell. 1992. A comparison of two scoring procedures with the NASA task load index in a simulated flight task. In *Proceedings of the IEEE National Aerospace and Electronics Conference (NAECON'92)*. IEEE, 734–740.

[48] Bilge Mutlu, Fumitaka Yamaoka, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2009. Nonverbal leakage in robots: communication of intentions through seemingly unintentional behavior. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction*. 69–76.

[49] Bastian Orthmann, Ilaria Torre, and Iolanda Leite. 2021. Auditory displays of robots' actions and intentions. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction Workshop on Sound in Human-Robot Interaction*.

[50] Stefania Pellegrinelli, Henny Admoni, Shervin Javdani, and Siddhartha Srinivasa. 2016. Human-robot shared workspace collaboration via hindsight optimization. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'16)*. 831–838.

[51] Elizabeth Phillips, Xuan Zhao, Daniel Ullman, and Bertram F Malle. 2018. What is human-like?: Decomposing robots' human-like appearance using the anthropomorphic robot (abot) database. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 105–113.

[52] Snejana Pleshkova-Bekiarska and Alexander Bekiarski. 2019. Building human mobile robot audio communication interface with artificial intelligence and deep learning. In *Proceedings of the International Conference on Creative Business for Smart and Sustainable Growth (CREBUS'19)*. IEEE, 1–5.

[53] Eric Rosen, David Whitney, Elizabeth Phillips, Gary Chien, James Tompkin, George Konidaris, and Stefanie Tellex. 2020. Communicating robot arm motion intent through mixed reality head-mounted displays. In *Robotics Research*. Springer, 301–316.

[54] Paul Saulnier, Ehud Sharlin, and Saul Greenberg. 2011. Exploring minimal nonverbal interruption in HRI. In *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'11)*. IEEE, 79–86.

[55] Matthias Scheutz, Bertram Malle, and Gordon Briggs. 2015. Towards morally sensitive action selection for autonomous social robots. In *Proceedings of the 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'15)*. IEEE, 492–497.

[56] Ivan Shindev, Yu Sun, Michael Coovert, Jenny Pavlova, and Tiffany Lee. 2012. Exploration of intention expression for robots. In *Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction*. 247–248.

[57] Moondeep C. Shrestha, Ayano Kobayashi, Tomoya Onishi, Hayato Yanagawa, Yuta Yokoyama, Erika Uno, Alexander Schmitz, Mitsuhiro Kamezaki, and Shigeki Sugano. 2016. Exploring the use of light and display indicators for communicating directional intent. In *Proceedings of the IEEE International Conference on Advanced Intelligent Mechatronics (AIM'16)*. IEEE, 1651–1656.

[58] Sichao Song and Seiji Yamada. 2018. Effect of expressive lights on human perception and interpretation of functional robot. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–6.

[59] Patric R. Spence, David Westerman, Chad Edwards, and Autumn Edwards. 2014. Welcoming our robot overlords: Initial expectations about interaction with a robot. *Commun. Res. Rep.* 31, 3 (2014), 272–280.

[60] Kyle Strabala, Min Kyung Lee, Anca Dragan, Jodi Forlizzi, and Siddhartha Srinivasa. 2012. Learning the communication of intent prior to physical collaboration. In *Proceedings of the IEEE 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'12)*. 968–973.

[61] Freek Stulp, Jonathan Grizou, Baptiste Busch, and Manuel Lopes. 2015. Facilitating intention prediction for humans by optimizing robot motions. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'15)*. IEEE, 1249–1255.

[62] Daniel Szafir, Bilge Mutlu, and Terrence Fong. 2014. Communication of intent in assistive free flyers. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*. 358–365.

[63] Daniel Szafir, Bilge Mutlu, and Terry Fong. 2015. Communicating directionality in flying robots. In *Proceedings of the 10th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI'15)*. 19–26.

[64] Pernilla Ulfvengren. 2003. *Design of Natural Warning Sounds in Human-Machine Systems*. Ph.D. Dissertation. KTH.

[65] Gentiane Venture and Dana Kulić. 2019. Robot expressive motions: A survey of generation and evaluation methods. *ACM Trans. Human-Robot Interact.* 8, 4 (2019), 1–17.

[66] Michael Walker, Hooman Hedayati, Jennifer Lee, and Daniel Szafir. 2018. Communicating robot motion intent with augmented reality. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 316–324.

[67] Frank Wallhoff, Jürgen Blume, Alexander Bannat, Wolfgang Rösel, Claus Lenz, and Alois Knoll. 2010. A skill-based approach towards hybrid assembly. *Adv. Eng. Inf.* 24, 3 (2010), 329–339.

[68] Dian Wang, Colin Kohler, Andreas ten Pas, Alexander Wilkinson, Maozhi Liu, Holly Yanco, and Robert Platt. 2019. Towards assistive robotic pick and place in open world environments. In *Proceedings of the International Symposium on Robotics Research*.

[69] Zhikun Wang, Katharina Mülling, Marc Peter Deisenroth, Heni Ben Amor, David Vogt, Bernhard Schölkopf, and Jan Peters. 2013. Probabilistic movement modeling for intention inference in human–robot interaction. *Int. J. Robot. Res.* 32, 7 (2013), 841–858.

[70] Samantha F. Warta, Olivia B. Newton, Jihye Song, Andrew Best, and Stephen M. Fiore. 2018. Effects of social cues on social signals in human-robot interaction during a hallway navigation task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 62. SAGE, Los Angeles, CA, 1128–1132.

[71] Atsushi Watanabe, Tetsushi Ikeda, Yoichi Morales, Kazuhiko Shinozawa, Takahiro Miyashita, and Norihiro Hagita. 2015. Communicating robotic navigational intentions. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'15)*. IEEE, 5763–5769.

[72] Tom Williams, Matthew Bussing, Sebastian Cabrol, Elizabeth Boyle, and Nhan Tran. 2019. Mixed reality deictic gesture for multi-modal robot communication. In *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI'19)*. IEEE, 191–201.

Received May 2020; revised April 2021; accepted May 2021