

Comparative Performance of Human and Mobile Robotic Assistants in Collaborative Fetch-and-Deliver Tasks

Vaibhav V. Unhelkar
Massachusetts Institute of
Technology
77 Massachusetts Avenue
Cambridge, MA, USA
unhelkar@mit.edu

Ho Chit Siu
Massachusetts Institute of
Technology
77 Massachusetts Avenue
Cambridge, MA, USA
hoseasiu@mit.edu

Julie A. Shah
Massachusetts Institute of
Technology
77 Massachusetts Avenue
Cambridge, MA, USA
julie_a_shah@csail.mit.edu

ABSTRACT

There is an emerging desire across manufacturing industries to deploy robots that support people in their manual work, rather than replace human workers. This paper explores one such opportunity, which is to field a mobile robotic assistant that travels between part carts and the automotive final assembly line, delivering tools and materials to the human workers. We compare the performance of a mobile robotic assistant to that of a human assistant to gain a better understanding of the factors that impact its effectiveness. Statistically significant differences emerge based on type of assistant, human or robot. Interaction times and idle times are statistically significantly higher for the robotic assistant than the human assistant. We report additional differences in participant's subjective response regarding team fluency, situational awareness, comfort and safety. Finally, we discuss how results from the experiment inform the design of a more effective assistant.

Categories and Subject Descriptors

H.1.2 [Models and Principles]: User/Machine Systems;
I.2.9 [Artificial Intelligence]: Robotics

General Terms

Experimentation, Performance, Human Factors

Keywords

mobile robotic assistant, human-robot collaboration, manufacturing, efficiency and productivity

1. INTRODUCTION

Industrial robots that operate in structured environments and perform repetitive tasks are widely used in manufacturing industries, especially those that are capable of building their factories around current industrial robotics

capabilities. For instance, approximately fifty percent of a typical automotive assembly process is performed using industrial robotics and automation [1]. However, many other industries still face significant challenges in automating low-volume, complex, and customized processes. Seventy percent of the build process for large commercial airplanes [2], ninety percent of the assembly process for electronics [3], and a majority of the work done by small business manufacturing in the US [4] is still manual. The automotive industry, too, sees challenges to further automation; final assembly of cars involves highly dexterous tasks that are performed almost exclusively by humans workers. Rather than supplant human workers with robotics, there is an emerging desire across these industries to strategically integrate advanced robotics into the manual processes. The goal is to improve efficiency by supporting people in doing their best work.

One approach is to use robotics to improve the productivity of people in performing “value-added work” and look for opportunities where robots can relatively easily perform “non-value-added work.” This paper explores one such opportunity, which is to field a mobile robotic assistant that travels between part carts and the automotive final assembly line, delivering tools and materials to the human workers. Human workers walk to pick their own parts today. Although seemingly insignificant, this walking time accumulates. An adept mobile robotic assistant can provide a significant boost to human worker productivity, while preserving operational flexibility as compared to fixed infrastructure solutions.

Multiple challenges must be addressed prior to deploying such a mobile robotic assistant in the dynamic, uncertain, and human-oriented assembly line environment [6, 7, 18]. The robotic assistant should not only be safe, but also perform the tasks efficiently and in a human-intuitive fashion. Akin to human-human collaboration, prior human-robot interaction studies have shown that a fluent collaboration requires awareness and anticipation of intent by both human and robotic agents [13]. To enable and facilitate this awareness and anticipation, it is important that both humans and robots communicate their status, location and intent, either explicitly through certain cues such as audio/visual signals or implicitly via their actions and motion paths [9]. Small deficiencies in the human-robot interaction in a time-critical domain such as automotive final assembly can significantly degrade the efficiency of the overall workflow.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
HRI'14, March 3–6, 2014, Bielefeld, Germany.
Copyright 2014 ACM 978-1-4503-2658-2/14/03 ...\$15.00.
<http://dx.doi.org/10.1145/2559636.2559655>.

In this paper, we evaluate the comparative performance of a mobile robotic assistant *vis-a-vis* human assistant during the delivery phase of a repetitive fetch-and-deliver task. Our goal is to gain a better understanding of the factors that impact the efficiency and effectiveness of a robot assistant situated in an analogue assembly line environment. We assess objective and subjective measures of team fluency, assistant saliency, and investigate the requirement and effectiveness of explicit and implicit indicators designed to improve the human co-worker’s awareness of the mobile robotic assistant. Specifically, we assess the effect of a flashing light on the robotic assistant and variations in the assistant’s approach angle. Lastly, we analyze whether a more *salient* assistant (utilizing the aforesaid indicators) produces a more *fluent* collaborator.

Results from the experiment indicate that interaction time and idle time are higher for the robotic assistant than the human assistant ($p < 0.05$). Changes in the assistant’s approach angle make the robot more salient according to subjective measures, but do not affect the objective measures of fluency. Further, although the explicit indicator (flashing light) is salient, it does not improve the objective measures of fluency and degrades the subjective measures and human situational awareness. We discuss how these results inform the design of a more effective assistant, and also suggest interesting questions regarding robot saliency and its affect on collaboration, which warrant further investigation and analysis.

2. RELATED WORK

Human-human collaboration during fetch-and-deliver tasks is seemingly trivial, and does not require much cognitive effort for the human agents. For instance, while delivering or receiving an object from a fellow human being we are least concerned about the motion of our hands. However, much can be learnt about how to successfully carry out human-robot collaboration by studying human-human interactions.

Prior studies on human-robot hand-overs provide useful insights for improving the motion of armed manipulators [8, 10, 17]. Experiments of give-and-take tasks between a human and robot [10], standing at a fixed locations, investigate the use of robot reaching gestures, vocalizations, and pauses as indicators for delivery. The study reports that communication using implicit natural actions is sufficient for give-and-take tasks, and the human collaborators do not require explicit instructions or expertise while collaborating with the robot. In other work, multiple studies on human-human and human-robot hand-overs result in design recommendations for robot motion, and formal descriptions of the physical and cognitive aspects of the hand-over process [17]. Experiments in [8] investigate the use of contrasting motion to communicate the onset of a hand-over, and demonstrate statistically significant improvements in fluency of hand-overs by using *temporal contrast* during delivery. The study also reports a small but not significant increase in robot waiting time when the human participant is performing an attention task. Although a mobile robot is used in [8], the primary focus is on the motion of the armed manipulator, and modifications to motion that reduce the human waiting time.

Interaction between *mobile* robots and humans has been investigated for fetch-and-carry tasks. Studies evaluate robot approach direction and its impact on task efficiency



Figure 1: The Rob@Work mobile robotic platform

and human comfort [14, 19]. Ideally results from these experiments would inform the design of motion planners that take into consideration human preferences. However, reported results are contradictory. [19] finds a left/right approach direction to be most favorable, while a rear approach direction to be the least favorable. In contrast, [14] indicates a frontal approach and delivery to be more favorable, especially for participants experienced with robots. These studies are carried out in settings without restrictions on approach directions, and with human participants that are primarily focused on the robot throughout its approach.

A frontal approach is often not practical in a constrained factory environment. Oblique approach angles affect the human’s visibility of the robot, and further study is needed into the effect of variations in robot approach angle from the rear direction. Additionally, human workers in a factory setting will usually be busy and not actively focused on the robot. Hence, there is a need to evaluate the robot’s approach towards an otherwise busy human co-worker.

3. AIM OF THE EXPERIMENT

Through human factors experimentation, we investigate interactions between an assistant and worker in an analogue assembly line environment, where the experiment participant takes the role of the worker. The task of the robotic or human assistant is to present the parts on a tray to the static human co-worker in a timely manner for continuation of the assembly task. Through the experiments we seek to evaluate the following hypotheses:

H1 *The interaction between the robotic assistant and worker during the delivery phase is less fluent than the interaction between a human assistant and worker. For this hypothesis, fluency is characterized by objective measures including interaction time and the assistant idle time.*

This hypothesis is founded in prior studies of hand-overs [17], which indicate hand-over quality degrades when working with a robot versus a human partner. We hypothesize that a similar effect exists, even when the assistant does not use manipulators and manipulation is the sole responsibility of the human worker. Idle times of agents (both human and robot) are indicative of fluency in a collaborative task, and have been used as objective measures of fluency in prior studies [13,

16, 15]. The design of our experiment ensures that the human worker is continuously occupied with tasks, and hence only idle times of the assistants (both robot and human) are evaluated. Interaction time, indicative of the total delivery time, is meant to quantify the time both the agents are interacting during the delivery phase of the task.

H2 *The worker subjectively assesses the interaction with the robotic assistant during the delivery phase to be less fluent than similar interaction with a human assistant. Subjective measures of fluency are evaluated using a series of Likert-scale questionnaires.*

Subjective measures of fluency are as important as their objective counterparts for evaluating human-robot collaboration. Hence, we evaluate the current hypothesis as a follow-on to **H1**. We have developed the questionnaire used in this experiment based on [12], which includes a survey of questions used to evaluate team fluency that produce values of Cronbach’s alpha ≥ 0.7 (indicating measurement of similar latent variable).

H3 *Salient indicators for the robotic assistant improve the worker’s awareness of the robot. Namely as indicators, we investigate the effect of variations in approach angle and the inclusion of a flashing light on the robot.*

Literature suggests awareness of the assistant and its intent improves task efficiency [13]. Factory settings are noisy, the workers’ attention is occupied with assembly tasks, and the robot may not always be in the human worker’s field of view. With this hypothesis we evaluate whether the specified indicators make the robot more salient. We measure the *notice time*, the time of the participant’s first head turn towards the assistant. An evaluation inspired by the Situational Awareness Global Assessment Technique (SAGAT) [11] is designed to measure the human worker’s awareness of the robot in the task environment.

H4 *Salient indicators for the robotic assistant improve the objective and subjective measures of fluency for the robot.*

With this last hypothesis we evaluate whether the indicators, which may make the robot more salient, do indeed influence task fluency. Namely, we investigate the effect of indicators described in **H3**. The objective and subjective measures described in **H1** and **H2** are used to assess improvement in task fluency.

4. EXPERIMENT METHODOLOGY

The experiment is designed to simulate an analogue environment to the assembly line. In this setting the participant worker is standing at a workstation, stationary, and facing away from the assistant’s approach path. The worker is provided an assembly task to occupy their attention. In the course of each trial two assistants, one robotic and one human, interact with the experiment participant. The assistants deliver parts enabling the continuity of the assembly task.

4.1 Materials and Setup

A Rob@Work robotic platform [5], shown in Fig. 1, is used as the mobile robotic assistant for this experiment. The

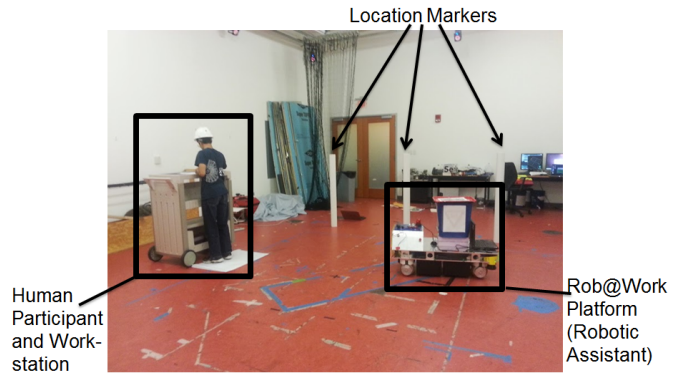


Figure 2: Experiment setup with human participant, workstation, Rob@Work platform, and white location markers.

robotic platform is designed for industrial environments and has an omni-directional base equipped with four independently actuated and steerable wheels, wheel encoders, two SICK S300 safety laser scanners¹, and safety features (both at hardware and software level) to avoid collisions. The basic platform is augmented with one Asus Xtion RGB-D device² used for person tracking, and one red, flashing and rotating light³. A raised platform and tray are mounted on the robot to represent the height of a future robotic arm. The robot, though capable of navigating autonomously, is operated manually by a human supervisor throughout the experiment to reduce variability in motion paths.

Observations during the experiment are obtained using demographic survey, four in-experiment questionnaires, a post-experiment questionnaire, a Vicon motion capture system, video camera, and the on-board Xtion. The participants and the human assistant wear hard hats with Vicon markers. The robotic assistant also has Vicon markers mounted on its top to enable tracking via the Vicon cameras.

During each experiment, a recording of factory sounds⁴ is played to simulate the factory environment. This serves to mask the noises made by the movement of the assistants to some degree, much like actual factory conditions would. Human participants are asked about the location of assistants in some questionnaires; thus, to eliminate the need to numerically estimate distances we place three large white poles equidistant from each other behind the participant. These poles are used as visual indicators to mark locations in the room (see Figs. 2-3), and divide the room into four names zones: *very close*, *close*, *far*, and *very far*, relative to the participant.

4.2 Procedure

Each participant is instructed to stand at a specified location to work at a standing table as shown in Fig. 2. A dexterous, model-assembly task is presented to the participant. The task is chosen for its complexity and similarity to a factory assembly task, and involves constructing a Lego model (Fig. 4). Portions of the Lego parts and assembly instructions are delivered at specified

¹<http://goo.gl/VwoA42>

²http://www.asus.com/Multimedia/Xtion_PRO_LIVE/

³<http://goo.gl/yxreko>

⁴<http://goo.gl/od11EQ>

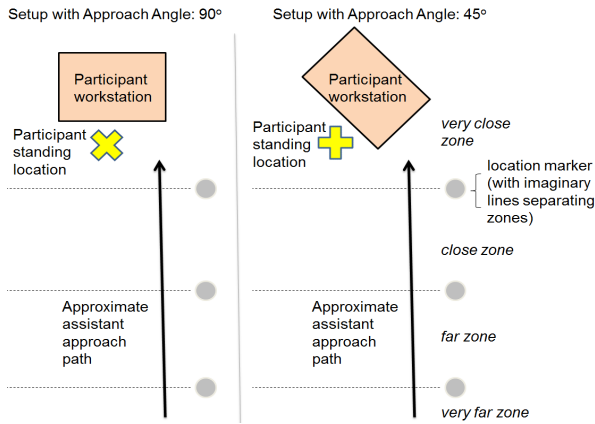


Figure 3: Schematic diagram of the experimental setup for the two approach angles (90° and 45°)

intervals during the experiment by either the human or robotic assistant. Each delivery, but for the first, consists of three items: Lego parts, corresponding instruction set, and a questionnaire to assess level of awareness of the robot, and perception of safety, comfort, trust, and fluency. To keep the participants occupied prior to the first delivery, the first set of instructions (but not the Lego parts and questionnaire) are given during the briefing, and the first delivery contains only Lego parts and a questionnaire. The total experiment task time is ≈ 12 -18 min, with a cumulative participant-assistant interaction time of ≈ 10 -20 seconds for each trial. This is similar to factory-like tasks, where delivery of parts constitutes only a portion of the overall assembly.

Each participant is aided by an assistant four times during the experiment, twice consecutively each by the human and robotic assistant (Fig. 5). This setup allows for each participant's response towards the robot to be compared to their own responses of working with a human assistant. Participants are randomly assigned to two groups one in which the first assistant is a human and the other in which the first assistant is a robot, to counterbalance differences that may arise due to the sequence of human and robot interactions. One of the experimenters plays the role of human assistant for consistency; the same human assistant works with all the participants. Both the human and robot assistants follow the same schedule for deliveries and stop at the same distance from the participant. Further, trays are used by both human and robotic assistants during delivery, since the focus of the study is on the assistant's approach rather than manipulator motion.

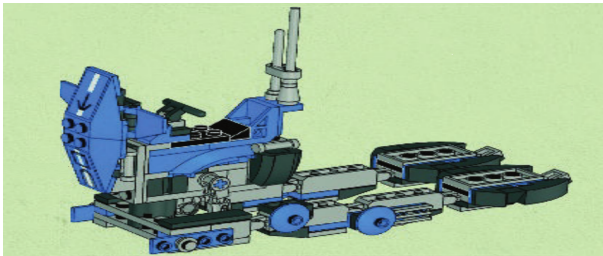


Figure 4: The Assembled Lego Model



Figure 5: Typical deliveries made by the Assistants (Top row: approach angle = 90° , Bottom row: approach angle = 45° , Left column: robotic assistant, Right column: human assistant)

Participants are instructed to remove all the items from the tray once an assistant stops next to them. Further, they are told before the experiment that the assistants are involved in other time-sensitive tasks, to indicate the importance of the assistant's time. Using the items delivered, the participants build the part of the Lego model corresponding to the current instruction, and then answer the corresponding questionnaire. The assistant's approach towards the human co-worker, for the next delivery, starts as soon as the participant begins answering the previously delivered questionnaire. The length of the questionnaire is designed to ensure the participant is always busy when the assistant (both human and robot) reaches him/her.

When a delivery is not in progress, the human and robot assistants either move around behind the participants in the *far* or *very far* zones or stand still. These movements are scheduled for pre-determined times during the experiment, and are included to simulate the assistant's involvement in other factory tasks. At different, pre-determined times the appearance of the robot is altered by affixing different colored foam bumpers and printed signs; similarly, the human assistant alters his appearance by wearing/removing a safety vest, watch, and gloves. The color of the trays used by both human and robotic assistants, on which items are delivered, also varies between deliveries. The participants are not informed of these signs, indicators, and possibility of their modification a priori, since our goal is to evaluate the saliency of these features in a natural workflow.

4.3 Design

The experiment is designed to evaluate the effect of three independent variables on the human-robot collaboration task, and participants are randomly assigned across a fourth factor to counterbalance differences that may arise due to sequence of human and robot interactions:

- **IV1:** Human or Robotic Assistant, varied within participants.
- **IV2:** Robot's Flashing Light (Explicit Indicator) On or Off, varied between participants.
- **IV3:** Assistant Approach Angle (Implicit Indicator) 90° or 45° , varied between participants.
- **IV4:** Human or Robotic Assistant for the first delivery, varied between participants.

Table 1: Design of Experiment (N=24)

		Approach Angle		<i>Participants</i>
		90°	45°	
Light	On	H-H-R-R (3)	H-H-R-R (3)	12
		R-R-H-H (3)	R-R-H-H (3)	
	Off	H-H-R-R (3)	H-H-R-R (3)	12
		R-R-H-H (3)	R-R-H-H (3)	
<i>Participants</i>		12	12	24

The approach angle variable indicates the orientation of participant relative to that of the assistant’s fixed approach path (Fig. 3). The 90° approach angle corresponds to the participant facing towards a wall and directly away from the assistant’s path. The 45° indicates an orientation of 45° away from the wall and towards the assistant’s path. We use light rather than sound as an explicit indicator since factory environments are frequently much too noisy for effective verbal/auditory communication.

Thus, the experiment is carried out as a mixed factorial design with four factors, one within- and three between-participants. Participants are randomly assigned across the 8 groups (2^k , where $k=3$ is the number of between-participant factors) indicated in Table 1. Next, we describe the dependent measures observed for each participant.

4.4 Objective Evaluation

Objective measures of team fluency include *interaction time* and *assistant idle time*. Interaction time is defined as the difference between the assistant’s stop on arrival and the beginning of the assistant’s retreat, and provides an assistant-centric measure of the time of each task’s delivery phase. Idle time is defined as the difference between the assistant’s stop on arrival and the start of the participant’s reach towards the tray. Idle time is a subset of the interaction time; we measure both since uncontrolled-for factors may affect the precise time of each assistant’s retreat. Saliency of the assistant is quantified by the derived measure *notice time*, the time between the participant’s head turn towards the assistant and the assistant’s arrival. Note that the *notice time* may be negative or positive depending on whether the assistant is acknowledged with a head turn prior to or after its arrival. Table 2 defines these measures for the delivery task in the current experiment. These

Table 2: Objective Measures: Definitions

		----- ----- ----- ----- -----			
		t_{notice}	t_{arrive}	t_{reach}	t_{retreat}
Interaction Time	$t_{\text{retreat}} - t_{\text{arrive}}$ Time between assistant’s stop on arrival and retreat, i.e., time required to complete the delivery				
Idle Time	$t_{\text{reach}} - t_{\text{arrive}}$ Time between assistant’s stop on arrival and the onset of the participant’s reach towards the tray.				
Notice Time	$t_{\text{arrive}} - t_{\text{notice}}$ Time between participant’s first head turn towards the assistant and the assistant’s stop on arrival.				

Table 3: Objective Measures - Question Set

Awareness of Assistant	
Common Questions	
1.	What color was the top surface of the tray on which the Lego pieces were delivered?
2.	After delivering the parts, what was the robot/human assistant doing while you were working on the model? (i.e. which zone was he/it in, was he/it stationary or moving?)
Robotic Assistant	
3.	What colors were the bumpers on the robot?
4.	What did the sign on the robot say the last time it delivered Legos?
5.	Were the lights on top of the robot on during the last delivery?
Human Assistant	
6.	What color was the assistant’s watch?
7.	What color were the assistant’s gloves?
8.	Was the human assistant wearing a safety vest?

measures are summed across the two deliveries, for each assistant, to obtain the cumulative measures reported in Section 5. The total task time in this study is dependent on the participant’s expertise in Lego assembly tasks. Hence, the total task time, a usual measure of task fluency, is not used in this study since the assistants only deliver the parts and do not contribute in the actual assembly.

These objective measures are independently coded by two raters from the video recordings of the experiment. The measured quantities are continuous and so Pearson’s correlation coefficient is used to determine the inter-coder agreement. For all the derived quantities, the resulting correlation coefficient is ≥ 0.98 indicating very high inter-coder agreement. Further, when determining statistical significance, data from both the coders is used independently to arrive at the final results, and any differences among coders are reported as errors.

To evaluate saliency of robot features, additional quantitative data is derived through in-experiment questionnaires. Participants are requested to fill out these in-experiment questionnaires between assembly steps. Questions are selected from a list of eight questions (in Table 3) pertaining to assistant’s whereabouts and characteristics, and responses are evaluated as *correct* or *incorrect*. The questions pertain to the time immediately preceding the delivery of the questionnaire, and the features, of both the robotic and human assistant, examined via these questions are hidden from the participant while he or she answers the questions. This evaluation is inspired by the Situational Awareness Global Assessment Technique (SAGAT) [11], and focuses on Level 1 SA pertaining to “the perception of the elements [including status, attributes, and dynamics] in the environment within a volume of time and space.”

4.5 Subjective Evaluation

Likert-scale statements and open-ended questions are used to collect subjective data regarding the experiment. Participants are administered a pre-experiment demographic survey and a post-experiment questionnaire. Further, subjective questions are also presented via the four in-experiment questionnaires. Several Likert-scale measures, derived from [12] and listed in Table 4, are used to evaluate participants’ subjective response regarding comfort, safety, and perceived fluency.

Table 4: Subjective Measures - Question Set

Comfort and Safety	
1.	I am comfortable with the time at which I first noticed the assistant.
2.	I feel safe working in this environment.
3.	I am comfortable with my level of awareness about the assistant’s whereabouts.
4.	I was stressed when the human/robotic assistant was assisting me.
5.	I felt safe when working with the human/robot assistant.
6.	I would have liked to notice the human/robot assistant coming earlier.
Fluency	
7.	The human/robotic assistant and I work well together.
8.	Deliveries made by the human/robotic assistant were smooth.
9.	I worked fluently together with the human/robot assistant.
Trust in Assistant	
10.	The human/robotic assistant’s actions were consistent.
11.	The human/robot assistant came when I expected him/it to.
12.	The human/robot assistant’s actions were predictable.
13.	The human/robot assistant was dependable.
Additional Indicators	
14.	The human/robot assistant did his/its part successfully.
15.	The human/robot assistant contributed to the success of the task.

The in-experiment questionnaire is changed from delivery to delivery. Only a subset of the questions shown in Tables 3-4 are provided in each questionnaire. This is to reduce learning effects.

5. RESULTS

This section summarizes statistically significant results, trends, and other insights obtained from the experiments.

5.1 Participants

Thirty participants performed the experiment, out of which data from six participants could not be used in analysis due to incorrect task performance, non-completion of the experiment, or missing video data. The results presented in this section are for the 24 participants that successfully completed the experiment with complete data for analysis. These participants were randomly assigned amongst the three between-participant factors of the experiment, resulting in three replicates for each experimental setting (Table 1). The participants (13 men and 11 women) were recruited via email, and had a median age of 20 years (max = 31, min = 18). None of the participants indicated any form of colorblindness in the pre-experiment demographic survey. Prior to the experiment a pilot study was carried out with 5 participants to streamline the experimental procedure; data from the pilot experiment are not included in the reported analysis.

5.2 Assistant Interaction and Idle times

A mixed factor, four-way Analysis of Variance (ANOVA) is carried out to compare the interaction times and assistant idle times across the four independent variables (1 within-, 3 between-participants). Statistically significant differences ($p < 0.05$, $N=24$) are observed in both these measures for use of human versus robotic assistant (the within-participant

variable). No statistically significant differences were found across the other factors including use of robot flashing light, approach angle, and type of first assistant.

Further analysis of the significant factors confirms that the interaction times and idle times associated with the robotic assistant are statistically significantly higher than those for the human assistant ($p < 0.05$, using two-tailed, paired t-tests to compare means with unknown, unequal covariance). The human assistant on average interacts with the participant for a cumulative time of 8.9 seconds ($SD = 3.1s$) as compared to the robotic assistant’s average of 12.7 seconds ($SD = 6.3s$). Similarly, the human assistant idles for a cumulative time of 1.0 second ($SD=2.4s$), in contrast to the robotic assistant’s idle time of 3.8 seconds ($SD = 6.0s$). These results support our first hypothesis **H1**, i.e., according to the objective measures of fluency the robotic assistant is a less fluent collaborator in comparison to the robotic assistant.

5.3 Subjective Measures of Fluency

Likert responses to statements about fluency (presented in Section 4.5) are analyzed across factors using non-parametric tests. No statistically significant differences are found in responses for the human versus robotic assistant (using the paired, Wilcoxon signed-rank test). Thus, the second hypothesis **H2** comparing the subjective measures of fluency is not supported by the experiment results.

5.4 Saliency

Saliency of the assistant is evaluated using the derived measure *notice time* as defined in Section 4.4. Responses to fourteen objective questions (listed in Table 3) pertaining to assistant’s whereabouts and characteristics are also evaluated and frequencies of *correct* versus *incorrect* answers are compared. Overall no statistically significant effects are found for type of indicators, and the third hypothesis **H3** is unsupported. Type or use of indicators did not produce statistically significant differences in interaction and idle times, leaving the fourth hypothesis **H4** unsupported as well.

However interesting differences emerge based on type of assistant, human or robot (within-participant variable). Analysis of notice time is performed using a mixed factor, four-way ANOVA, and a statistically significant difference ($p=0.005 \pm 0.001$) is observed. The robotic assistant is noticed much earlier with average notice time of 9.0 seconds ($SD=9.7s$) prior to stop on arrival, as opposed to the human assistant’s average notice time of 1.8 seconds ($SD=3.3s$). These results suggest that degradations in fluency are likely due to factors other than robot saliency, since on average the robot is acknowledged earlier than the human assistant.

Overall, there is only one statistically significant difference in the frequencies of correct responses between types of assistants. Color of the tray is noticed significantly better by the participants during the two deliveries each made by the robotic assistant (37%) than the human assistant (8%) ($p < 0.05$, Fisher’s exact test with 2×2 contingency table). However, participants demonstrated through responses on Question 2 (Table 3) that they were significantly more aware of their background environment after delivery by a human assistant ($p < 0.001$). This suggests that the robot may have a transitory distracting effect that degrades situational awareness, even after the robot leaves the participant’s side.

Table 5: Sample of Open-ended Comments

Human Assistant
“I liked that he said thank you!”
“Making eye contact & speaking was key to feeling more comfortable - I liked getting confirmation that the person wanted me to take stuff from the tray.”
“Delivered parts when I’m still working, made me feel more stressed.”
Robotic Assistant
“Smooth transition. Didn’t get too close to me which I liked.” “Did a good job at a simple task.”
“Having the robot moving around in the background was more distracting than the human in the background.”
“With the robot, I think I made it wait till I’m done to get the stuff, I was less stressed.”

More visible features such as safety vest for human assistant and the state of the light for the robotic assistant were noticed by 79% and 67% of the participants, respectively. The participants were not informed about the existence or relevance of these signs in advance, and thus noticed them during the course of natural interaction with the assistant. Participants were equally unaware about less noticeable features such as the human assistant’s glove color (13% correct responses) and watch (18%), and the robot’s bumper color (4%) and printed signs (21%).

5.5 Factors affecting Subjective Measures

The effect of flashing light and approach angle (between-participant factors) on Likert statement responses is evaluated using the two-sided unpaired Wilcoxon rank sum test. Participants agreed less strongly with the following statements when the robot’s light was flashing, indicating a reduction in perception of safety and trust in the robot:

- I felt safe while working with the robot. ($p < 0.05$, evaluated during post-experiment survey)
- The robot assistant’s actions were consistent. ($p < 0.01$ through in-experiment questionnaire, while not statistically significant during post-experiment survey)

Variation in approach angle resulted in statistically significant differences in responses as well. Participants agreed more strongly with the following statement in the 45° approach condition, indicating increased comfort with the robot when it approaches obliquely as opposed to from the rear (no such difference was found for the human assistant):

- I am comfortable with my level of awareness about the robot assistant’s whereabouts. ($p < 0.05$, assessed twice in the in-experiment questionnaires)

5.6 Open-ended Comments

The participants are asked to provide open-ended comments about their experience with each assistant at the end of the experiment. Selected comments from these responses are included in Table 5. The open-ended responses suggest mixed reactions towards the robotic assistant. Interestingly, some comments reflect that participants felt rushed by the human assistant, and that the robot let them work at their own pace.

6. DISCUSSION

The results of the human factors experimentation support our hypothesis (**H1**) that the human-human collaboration is more fluent, as quantified by assistant interaction times and idle times, than human-robot collaboration for the current fetch-and-deliver task. Statistically significant results indicate that the robotic assistant spent on average 3.8 seconds (43%) more time than the human assistant interacting with the human for the same delivery tasks. Similarly, a statistically significant three-fold rise (2.9 seconds, 3x) is observed in the assistant idle time. A factory assistant will be making deliveries to multiple human co-workers. These differences in interaction and idle time, though small in magnitude, will substantially affect the productivity of the robotic assistant over the course of a two or three-shift work day, and should be alleviated using design interventions. For example, a back-of-the-envelope analysis indicates if the experiment task was replicated over a two-shift day (assuming a total of 12 hr of assembly time), and the robot assistant is multi-tasked to deliver items to five workers, the robot accumulates approximately 30 minutes of *additional* interaction time, as compared to a human assistant.

Surprisingly, the robotic assistant’s approach is noticed on average much earlier (7.2 seconds, a statistically significant difference) by the participants as compared to that of the human assistant. Nonetheless, the robotic assistant idled more than the human assistant. This provides contrary evidence for the fourth hypothesis (**H4**), suggesting a salient agent does not necessarily produce an efficient collaborator.

These results suggest that degradations in fluency are likely due to factors other than robot saliency, since on average the robot is acknowledged earlier than the human assistant. While the participants reported the human and robotic assistant to be equally fluent, they appeared to be more comfortable with making the robot wait. This is supported by open-ended responses indicating different attitudes towards the human and robot. We posit that the human assistant’s time is valued more than that of the robot, and personal objectives and comfort take a higher priority during collaboration with the robotic assistant. Further study is required to understand how to design the robot and its human interface so that it does not wait unnecessarily for human co-workers, which degrades the productivity of the robotic assistant and the overall assembly line workflow.

We confirm that an oblique approach angle is preferable since the participants report increased comfort about their awareness of the robot, although it did not improve the objective measures of fluency. The red flashing light is observed to be the most noticeable feature amongst those evaluated but did not improve objective measures of fluency. Further, the participants reported feeling less safe with the light flashing. This is possibly due to the color choice, and suggests that a red flashing light should not be used in nominal, safe operation of the robot.

7. LIMITATIONS AND FUTURE WORK

We designed our experiment to emulate a factory setting through careful choice of task and features such as noise, however limitations remain. The study was carried out in a large mostly empty room with student participants rather than factory workers. Further, the participants were working with the robotic assistant for the first time, and

hence the effect of long-term experience working with the robot cannot be evaluated. In typical factory operations, the robot will be assisting multiple workers. Although participants were told that the assistant has additional responsibilities and tasks, it is possible human workers will behave differently when the robot's responsibilities clearly relate to other co-workers. Future experiments will include a longitudinal study with multiple co-workers, to observe how the interaction changes over time and with multiple people. We also plan to conduct user studies in a factory setting to investigate whether the results, trends and insights from this study apply in the real-world operational environment.

8. CONCLUSIONS

We conduct a human factors experiment to compare the performance of a mobile robotic assistant and human assistant during the delivery phase of a fetch-and-deliver task. Results from the experiment indicate that interaction times and idle times are statistically significantly higher for the robotic assistant than the human assistant. However the robotic assistant's approach is noticed on average much earlier by the participants as compared to that of the human assistant. These results suggest that degradations in fluency are likely due to factors other than robot saliency. Based on our observations, we conjecture that the human assistant's time is valued more than that of the robot, and personal objectives and comfort take a higher priority during collaboration with the robotic assistant. We confirm that an oblique approach angle is preferable since the participants report increased comfort about their awareness of the robot. The robot's red flashing light did not improve objective measures of fluency, and the participants reported feeling less safe with the light flashing. This suggests that a red flashing light should not be used in nominal, safe operation of the robot.

9. ACKNOWLEDGMENTS

This work is supported by BMW. We thank our colleagues Stefan Bartscher and Johannes Bix of the Innovation Production Division, BMW Group, Munich, Germany for their valuable insights and domain expertise that informed this work.

10. REFERENCES

- [1] Stefan Bartscher, Innovation Product Division, BMW Group, personal communication, February 2012.
- [2] David Amirehteshami, Boeing Research and Technology, personal communication, July 2011.
- [3] John Davidson, Advanced Technology Engineering, Jabil, personal communication, June 2012.
- [4] Robots for small business: A growing trend, robotics online, oct. 31, 2003. http://www.robotics.org/content-detail.cfm/Industrial-Robotics-News/Robots-For-Small-Business:-A-Growing-Trend/content_id/1118.
- [5] Rob@work 3. <http://www.care-o-bot-research.org/robotwork-3>.
- [6] J. Bix. Mobile robots at the assembly line. In *International Conference on Robotics and Automation (ICRA), 1st Workshop on Industrial Mobile Assistance Robots*, 2013.
- [7] J. C. Boerkoel Jr and J. A. Shah. Planning for flexible human-robot co-navigation in dynamic manufacturing environments. In *International Conference on Human-Robot Interaction (HRI), Proceedings of Human Robot Interaction Pioneers*, 2013.
- [8] M. Cakmak, S. S. Srinivasa, M. K. Lee, S. Kiesler, and J. Forlizzi. Using spatial and temporal contrast for fluent robot-human hand-overs. In *Proceedings of the International Conference on Human-Robot Interaction (HRI)*, pages 489–496, 2011.
- [9] A. D. Dragan, K. C. Lee, and S. S. Srinivasa. Legibility and predictability of robot motion. In *Proceedings of the International Conference on Human-Robot Interaction (HRI)*, pages 301–308, 2013.
- [10] A. Edsinger and C. C. Kemp. Human-robot interaction for cooperative manipulation: Handing objects to one another. In *Proceedings of the International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2007.
- [11] M. R. Endsley. *Situation awareness analysis and measurement*, chapter Direct measurement of situation awareness: Validity and use of “SAGAT”. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
- [12] G. Hoffman. Evaluating fluency in human-robot collaboration. In *International Conference on Human-Robot Interaction (HRI), Workshop on Human Robot Collaboration*, 2013.
- [13] G. Hoffman and C. Breazeal. Effects of anticipatory action on human-robot teamwork: Efficiency, fluency, and perception of team. In *Proceedings of the International Conference on Human-Robot Interaction (HRI)*, pages 1–8, 2007.
- [14] K. L. Koay, E. A. Sisbot, D. S. Syrdal, M. L. Walters, K. Dautenhahn, and R. Alami. Exploratory study of a robot approaching a person in the context of handing over an object. In *AAAI Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics*, 2007.
- [15] S. Nikolaidis and J. Shah. Human-robot cross-training: computational formulation, modeling and evaluation of a human team training strategy. In *Proceedings of the International Conference on Human-Robot Interaction (HRI)*, pages 33–40, 2013.
- [16] J. Shah, J. Wiken, B. Williams, and C. Breazeal. Improved human-robot team performance using chaski, a human-inspired plan execution system. In *Proceedings of the International Conference on Human-Robot Interaction (HRI)*, pages 29–36, 2011.
- [17] K. W. Strabala, M. K. Lee, A. D. Dragan, J. L. Forlizzi, S. Srinivasa, M. Cakmak, and V. Micelli. Towards seamless human-robot handovers. *Journal of Human-Robot Interaction*, 2(1):112–132, 2013.
- [18] V. V. Unhelkar, J. Perez, J. C. Boerkoel Jr, J. Bix, S. Bartscher, and J. A. Shah. Towards control and sensing for an autonomous mobile robotic assistant navigating assembly lines. In *International Conference on Robotics and Automation (ICRA)*, 2014. Currently under review.
- [19] M. L. Walters, K. Dautenhahn, S. N. Woods, and K. L. Koay. Robotic etiquette: results from user studies involving a fetch and carry task. In *Proceedings of the International Conference on Human-Robot Interaction (HRI)*, pages 317–324, 2007.