Relative importance of spatial and temporal precision for user satisfaction in human-robot object handover interactions

Ansgar Koene¹, Anthony Remazeilles², Miguel Prada³, Ainara Garzo², Mildred Puerto², Satoshi Endo¹ and Alan M. Wing¹

Abstract. In current society there is a growing call for robotic platforms designed to provide direct assistance to humans within the near future. An essential requirement, if robots are to become accepted as service providers that interact directly with humans, is that the human-robot interactions must not only be safe and reliable, but also produce a satisfying user experience. Here we report results from a study on human-robot object handovers in which we show that timing aspect of the robot response outweighed spatial precision for determining user experience ratings.

1 INTRODUCTION

Safe, reliably satisfying and fluent object handover interaction between robots and humans has been identified as a key competence that is required if robots are to make the transition from assembly line machines to household and small business service robots (e.g. [1,2,3]).

Human-Robot interaction poses many challenges for the development of service robots. Human behaviour, though perhaps globally predictable is notoriously variable at the level of individual interactions. The design of efficient and autonomous service robot gets even more challenging when considering the great variability of surrounding environment conditions and, above all, the justifiably high premium placed on maintaining the safety of the user. As a result handover interactions between robots and humans tend to tend to be precise but slow, to the extent of sorely testing the patience of observers and any potential users (e.g. as illustrated by this example of a Human-Robot interaction from the LAAS¹, France [4]).

One aspect of human-human handover behaviours that until recently received very little attention in studies of reaching movements [5], and is thus poorly reproduced in human-robot interaction works, is the dynamic nature by which both participants adjust their movements according to the behaviour of their interaction partner. Recent studies on joint action between humans, including object handover, highlighted the rapid adaptation of interaction partners over repeated interactions, leading to reduced action delays as the partners become better able to anticipate the handover location [6,7].

In general, human behaviour frequently demonstrates speed-accuracy trade-offs in which actors exhibit a willingness to sacrifice interaction accuracy for higher speed [8,9]. Nevertheless, this does not guarantee that users would have the same attitude towards speed-accuracy trade-offs with robotic partners. The study presented in this paper therefore investigated if future users of service robots would accept this compromise during human robot interaction.

In this paper we report some of the results of the Coglaboration project (http://www.coglaboration.eu) in which we are developing a service robot that is optimised for fluent Human-Robot object handover interactions based on the behavioural study of the corresponding Human-Human interactions. The analysis performed in this document results from an experimental study in which various persons were exchanging objects with an autonomous robotic arm. In this experimentation, different robot control strategies as well as different human postures where considered to understand their impact on the user perception of the exchange quality. The interactions were evaluated combining qualitative response ratings, for each object exchange between the human and the robot, with motion tracking data of both partners. This combined data set was analyzed to operationalize the qualitative experiences of the participants, e.g. what had more impact on the subjective experience of the participants, the response speed of the robot or the accuracy of the robot movement? Answering such questions about the subjective priorities assigned to the speed and accuracy of a robotic interaction partner may provide important guidance for further developments in service robotics.

The paper is organized as follows. First, we describe the experiments, including the design, procedure, setup and robotic system we used (Section 2). Next, we describe the analysis of the data and its results (Section 3). In Section 4, we present our conclusions and discuss the implications for the development of robots that are designed for interacting with humans.

2 EXPERIMENTS

This section describes the experiment design, procedure, equipment and participants that were used to collect the data for our analysis.

¹ Dept. of Psychology, Univ. of Birmingham, B15 2TT, UK. Email: {a.r.koene, a.m.wing, s.endo}@bham.ac.uk
² Assistive Technologies Department of the Health Division of Tecnalia Research and Innovation, Parque Tecnologico de San Sebastian, E-20009 San Sebastian, Spain. Email: {name.surname}@tecnalia.com.
³ https://www.youtube.com/watch?v=gK0e9zYQXY
2.1 Experiment design

The design of the experiment was inspired by the small-business scenario of a robotics assistant to a car mechanic, considering that a robotic assistance could have a significant impact in similar working areas. The emulation of the industrial setting was mainly intended to force the person to maintain different posture, with the objective of characterising how these postures may affect the perception of the exchange quality. We associated three different configurations to this scenario, related to three different human postures, referred to as ‘Engine Bay’, Hydraulic Lift and ‘Lying under the car’:

- The ‘Engine Bay’ configuration simulates working on the engine of a car with the person bent over task area, in contact with the simulated engine, able to view the object handover; generally reaching to the right and slightly backwards with slight movement range restriction; in that configuration, the robot has a full view of the person (Figure 1a).
- The ‘Hydraulic Lift’ configuration simulates working under a car on a hydraulic lift with the person reaching slightly above the head while maintaining contact with the car; free to observe the object handover; generally reaching to the right side with range of movement only slightly impaired by keeping contact with the car using one hand; the robot has a full view of the person (Figure 1b).
- The ‘Lying under the car’ configuration simulates lying under the car with limited view; lying on back with a limited range of arm movement; the robot has a limited view of the person (Figure 1c).

The task for the participants was to receive an object (a flashlight) from the robot, take the object to the task area (e.g. ‘Engine bay’) and then hand it back to the robot. Thus each trial consisted of two phases:
1. a robot-to-human handover (R→H)
2. a human-to-robot handover (H→R).

2.2 Experiment procedure

The experiments were conducted over three days, so that a single scenario configuration was tested per day, avoiding the need for frequent reconfigurations of the setup between trials and minimizing participant fatigue and reducing habituation. Some habituation effects however remained, primarily with respect to the orientation with which participants would expect the object to be presented during the handover.

Each configuration was run with 7 naïve participants. The participants (4 male, 3 female) were recruited from Tecnalia staff, 5 of whom had little or no prior experience of interacting with robots.

To test the effect of speed-accuracy trade-off on the user experience the robot movements were executed at 5 different speeds such that the robot movement was approximately 1, 1/2, 2/5, 1/3 or 1/4 as fast the average human movement (0.55m/s) for performing the handover reach.

The experiment started with five object handovers at ‘normal’ speed (approximately 0.27m/s, i.e. ½ human speed), permitting the participant to become familiarized with the robotic system and the evaluation protocol. All other trials were then performed in a random order with each of the 5 speeds repeated 3 times. In total each participant performed (5 speeds x 3 repeats + 5 practice) x 2 directions x 3 configurations = 120 trials, giving a total of 120 x 7 participants = 840 trials.

Each trial consisted of:
1. Exchange from robot to human (R→H)
   1.1. Human requests object by reaching towards the robot.
   1.2. Robot brings the object to the human.
   1.3. Human takes the object from the robot and brings it to the task area, i.e. the ‘car’.
   1.4. The participant evaluates the handover
2. Exchange from human to robot (H→R)
   2.1. Human holds the object out towards the robot.
   2.2. Robot reaches for the object.
   2.3. Robot takes the object back to itself.
   2.4. Human evaluates the handover.

Once all the trials were completed, the participant was interviewed to provide additional feedback and qualitative evaluation of their experience during the interactions.

The evaluation ratings were provided by the participant after each handover using a touch screen. The evaluation statements were:

- Q1: It was easy to receive the object
- Q2: I was satisfied with the interaction
- Q3: The interaction was comfortable
- Q4: I felt safe during the interaction

These statements were chosen to gauge the user-friendliness of the interaction experience. Q1 checked if the person felt any difficulty to make the exchange, to test the compliance of the robot behaviour to natural human interactions. Q2 is similar to Q1 but puts the rating in the context of the environment (or object) properties that might reduce the ease of the handover, e.g. the ‘Lying under the Car’ configuration even a highly satisfying interaction might not be easy, due to the posture. Q3 on the other hand is highly dependent on environmental factors and thus reflects the degree to which the robot managed to ameliorate these. Q4 finally explores the perceived threat of the robot and its behaviour. For each of the questions the subjects were asked to enter an evaluation score between 1 (fully disagree) and 9 (fully agree).

In addition to the qualitative evaluation, a set of quantitative data was recorded during the experiments, including:

- The location of the human hand as a function of time, providing movement kinematics.
- The articular pose of the robot as well as the measured efforts per joint.

Figure 1. Human postures in the three car mechanic configurations
• The timing of the events during the exchange procedure (robot motion start, end, contact trigger, hand manipulations)

2.3 Experiment setup & equipment

The car mechanic scenario was simulated using the frame shown in Figure 2. The flat surface (dark shaded area in the design drawing on the left) defined the task area for the mechanic, i.e. the area that participants were required to bring the object to. The task surface was moved from waist height, for the ‘Engine Bay’ configuration, to above the participant’s head, for the ‘Hydraulic Lift’ configuration. For the ‘Lying under the car’ configuration the participants were required to lie on a height adjusted bed, to accommodate the workspace limitations of the robotic system.

The work frame (and the bed in the “Lying under the car” configuration) was positioned such that the participant was at a distance of 100-125cm from the robot. This allowed the participant outstretched hand to be inside the robot’s workspace while his/her body remained safely beyond reach of the robot. The qualitative evaluation was performed captured through a laptop with touch-screen.

Concurrently, the following equipment was used to collect quantitative data. A Polhemus Liberty magnetic motion tracking system (Polhemus Inc., Vermont, USA) with four magnetic markers, enabled to record the position and orientation of the participant’s arm at 240Hz during the object handover. The magnetic markers were placed onto the participant’s right arm (one on the shoulder, one on the back of the hand, one on the thumb and one on the index finger).

2.3.1 The Robotic system

The robotic system used is composed of the KUKA’s lightweight robot (LWR)[10] mounted on a frame, as illustrated on Figure 3a onto which the Prensilia’s hand [11] is mounted at the end-effector (Figure 3b).

The complete setup with the robotic system, work frame and participant positioning is illustrated in Figure 4 (left image: ‘Lying under the car’ configuration; right image: ‘Hydraulic Lift’ configuration).

The perception was performed through a Kinect camera (at

![Design drawing of work frame](image1)

![Mechanic Scenario work frame](image2)

![Bed for the ‘Lying under the car’ configuration](image3)

**Figure 2.** Work frame and bed used for simulating the ‘Car Mechanic’ scenario

![Figure 3. a) KUKA’s LWR; b) Prensilia’s hand](image4)

**Figure 3.** Illustration of the setup and robotic system.

The control law guiding the robot motion depending on the visual perception is based on the DMP formalism (*Dynamic Movement Primitives*) [12,13]. This control approach permits to reproduce a reference motion pattern while maintaining a reactive convergence towards the possibly changing targeted location (the human hand for performing the exchange in our case). This framework was specialized in the context of the Coglaboration project to reproduce a human-like motion, using as reference pattern a human arm motion extracted from human behaviour analysis. This shape driven mechanism is completed in the DMP approach with a feedback mechanism that permits to ensure a convergence towards the goal (human hand location). The overall control progressively switches from the shape driven behaviour (reproducing the learned trajectory) to the goal driven behaviour (to converge towards the hand of the person). Since the robot control mechanism is not the focus of this paper, we would like to redirect the interested readers to the two previously cited papers that respectively describe the DMP specialization for human robot object handover [12], and an initial validation of the control law [13] performed through comparison with a human motion database created in the project (see the ‘CogLab, Human Object Handover’ entry at the University of Birmingham, Behaviorinformatics wiki [14]).
3 DATA ANALYSIS AND RESULTS

In order to increase the statistical power of the analysis the data of each of the participants was combined to test for group level effect. In order to compensate for individual differences in mean (and/or range of) responses/behaviours the participants’ data was first normalized, with respect to the mean and inter-quartile ranges of their respective qualitative responses and quantitative measures. Data from the various participants of a given configuration were then pooled to increase the number of samples per manipulation from 3 per each of the 5 ‘Speeds’ to 21 samples per speed respectively.

3.1 Robot movement speed effect on perceived handover quality

Preliminary analysis of the rating responses by the participants showed a skewed response distribution with more than 60% of ratings given as 8 or 9 and less than 10% of ratings at 5 or less (Figure 5). These generally high ratings suggest that participants generally considered the interactions as satisfactory and safe. The skewed distribution of responses however violates the basic assumptions of parametric statistical tests (i.e. normal-distributed data), thus dictating that the subsequent statistical analyses had to be performed using non-parametric tests.

Comparison between the histograms of the five robot movement speeds (see colour coding) revealed that for the faster movement trials (Speed 1 = same speed as human) Comfort, Satisfaction and Ease of handover had more ‘9’ ratings than the slower movement trials (Speed 5 = ¼ speed of a human). As one might expect, however, for the ‘Safe’ rating this pattern was reversed, with more ‘9’ rated trials for the slower movement trials. Slow movements were thus considered as more safe. The participants’ preferences, however, nevertheless tended to favour the handover interactions in which the robot reacted quicker, suggesting that the key is to find a good tradeoff between these two participant desires (comfort, satisfaction, ease vs. safety).

Next we ran a Kruskal-Wallis non-parametric ANOVA [15] test to confirm that the apparent effect of the robot movement speed manipulation was indeed statistically significant (p<0.05). The results of the analysis are summarized in Table 1 for each experiment configuration, handover direction and qualitative rating category.

Table 1. Kruskal-Wallis non-parametric one-way ANOVA results to identify an effect of robot movement speed condition on qualitative rating responses. (p<0.05 indicates a statistically significant effect)

<table>
<thead>
<tr>
<th>Rating questions</th>
<th>Engine Bay</th>
<th>H. Lift</th>
<th>Lying u. C.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy</td>
<td>R ➔ H</td>
<td>R ➔ H</td>
<td>R ➔ H</td>
</tr>
<tr>
<td></td>
<td>.015</td>
<td>.000</td>
<td>.028</td>
</tr>
<tr>
<td>Satisfied</td>
<td>.205</td>
<td>.002</td>
<td>.504</td>
</tr>
<tr>
<td>Comfortable</td>
<td>.001</td>
<td>.000</td>
<td>.469</td>
</tr>
<tr>
<td>Safe</td>
<td>.000</td>
<td>.654</td>
<td>.238</td>
</tr>
</tbody>
</table>

In addition to confirming that the speed manipulation affected the participants’ Ease, Satisfaction, Comfort and Safety experience (p<0.05 for 3 out of the 4 questions in both the R➔H and H➔R directions for ‘Engine Bay’ and ‘H. Lift’; grey shaded cells), we note that the ‘Lying under the car’ configuration showed no statistically significant effect (p>0.05). A possible reason for this may have been the reduced ability to see the robot motion and the reduction in the speed of the human movements in this posture. This illustrates the complex nature of the perceived exchange quality, which was affected by more variables than just the robot velocity.

3.2 Correlation of qualitative ratings with quantitative handover movement properties

Next we analysed the correlation between the qualitative ratings (captured by the touch-screen laptop) and the quantitative measures of the human (captured by the Liberty magnetic motion tracking system) and robot (captured by the configuration log of the robotic arm) movements in order to identify more clearly if the changes in perceived Human-Robot interaction quality were primarily related to timing aspects of the movement or spatial accuracy related aspects.

The quantitative measures we considered were:

1. Two properties relating to temporal aspects of the behaviour.
   1.1. End time difference, i.e. end time of the robot reaching movement – end time of the human reaching, which is directly related to the time the human had to wait for the robot in order to complete the handover.
   1.2. Peak velocity time difference, i.e. time when the robot arm reached peak velocity – time when the human arm movement reached peak velocity. This measure was included in addition to 1.1 to reflect the shape of movement profiles and also for the psychological impact of the moment of peak velocity.

2. Two properties relating to the spatial aspects of the behaviour.
   2.1. Movement distance difference, i.e. the distance between start- and end-point of the human reaching movement – the distance between the start- and end-point of the robot reach. This measure provides a sense of the relative amount of ‘effort’ the person
would consider that each partner was contributing to the handover interaction.

2.2. End position error, i.e. the distance between the robot movement end-point and the human reach end-point (before the human made an optional final adjustment for picking up/placing the object from/in the robot hand if needed). This provides a direct indication of the spatial accuracy of the robot performance.

3. One spatio-temporal property.

3.1. Robot peak-velocity, a factor with clear psychological links to perceived threat/aggression (when high) or disinterest/reluctance (when low).

The Spearman rank-correlation⁴ [16] was computed between each of the four qualitative ratings (‘Easy’, ‘Satisfied’, ‘Comfortable’ and ‘Safe’) and the five quantitative measures for both $R \rightarrow H$ handover and the $H \rightarrow R$. The results are summarized in Figure 6. The three subplots provide the results for the ‘Engine Bay’ (top), ‘Hydraulic Lift’ (middle) and ‘Lying under the car’ (bottom) configurations respectively. The x-axis lists the quantitative measures (e.g. ‘End time difference’) for the $R \rightarrow H$ and $H \rightarrow R$ handover phases, grouped by measures relating to ‘timing performance’ on the left (first 4) and measures relating to ‘spatial performance’ on the right (last 4). The y-axis lists the four qualitative rating. Colours indicate the value of the Spearman rank-correlation coefficient ($\rho$), with warm (red) colours for positive correlations and cold (blue) colours for negative correlations (see legend on right). Correlations of $|\rho| < 0.2$ are generally not significantly different from zero ($p > 0.1$).

The main observations from the correlation analysis are summarised below.

The Safe rating and the Easy-Comfortable-Satisfied ratings have generally opposing relationships to the quantitative behaviour properties, i.e. participants were more satisfied when the interaction was fast and the robot reached close to their hand but at the same time felt more safe when the interaction was slow and the robot’s reach ended further away from them.

Surprisingly, the negative correlation between ‘Robot peak velocity’ and perceived ‘Safety’ did not reach significant levels ($p < 0.2$) even though ‘Robot peak velocity’ was strongly correlated with ‘Ease’, ‘Satisfaction’ and ‘Comfort’. This suggests that different parts of the robot velocity profile (e.g. early acceleration, final deceleration) may have different degrees of impact on the perceived handover quality.

The ‘timing performance’ related measures were more strongly correlated with the qualitative ratings while the ‘spatial performance’ measured were mostly weakly correlated to the ratings. This indicates a willingness by the participants to engage in a speed-accuracy trade-off in favour of rapid interactions.

Overall, the strongest correlations between qualitative ratings and objectively observable quantitative measures was found in the ‘Hydraulic Lift’ configuration, though the ‘Safe’ ratings also showed strong correlations in the ‘Engine Bay’ configuration. ‘Lying on the Bed’ yielded the weakest correlations, probably because the reduced visibility and reduced freedom of movement in this configuration dominate the participant’s experiences.

‘End position error’ was found to have its strongest (negative) correlations in the ‘Lying under the car’ configuration, suggesting that when considering the speed-accuracy trade-off for Human-Robot interaction, users may prefer a greater weight on speed, when they themselves have the space to make compensation movements to adapt to robot errors, while they prefer to wait for more accurate robot behaviour when they themselves are constrained in their movement space, as was the case in the ‘Lying under the car’ configuration.

4 CONCLUSIONS AND IMPLICATIONS

In this paper we reported on a series of Human-Robot object handover interaction experiments in which we manipulated the posture of the human (using three configurations: ‘Engine Bay’, ‘Hydraulic Lift’ and ‘Lying under the car’) and the speed of the robot reaching movement while gathering qualitative rating responses and quantitative movement recordings for each handover.

4.1 Conclusions

Analysis of the data showed that participant rating responses were more strongly correlated to temporal aspects of the robot movement than they were to the spatial performance.

One of the strongest determinants for the participant’s ‘Satisfaction’, ‘Ease’ and ‘Comfort’ ratings was the duration it took for the robot to reach the handover location. This was reflected in:

1. the negative correlations with the ‘timing performance’ related measures showing that the ratings decreased whenever the time difference between the human and robot movement end, or peak velocity time, increased (Figure 6, ‘Hydraulic Lift’), possibly because in everyday human-human interactions such delayed responsiveness is often associated with a reluctance to cooperate;
2. post-session interview statements by participants in favour of the faster speeds and disliking the slow speeds.
There is an inherent tension between the participants’ perceived ‘Ease-Satisfaction-Comfort’ and perceived ‘Safety’, as illustrated by the finding that the ‘Safe’ rating and the ‘Easy-Comfortable-Satisfied’ ratings generally have opposing relationships to the quantitative behaviour properties (Figure 6, opposite sign of the correlation coefficient).

Somewhat surprisingly, while the perceived ‘Safety’ had a strong positive correlation ($\rho > 0.4$ in ‘Engine Bay’) with measures relating to the relative timings of the Human and Robot movements (‘End time difference’, ‘difference in time of Peak velocity’) it was only weakly (negatively) correlated with direct measures of the robot kinematics (‘Robot peak velocity’) (Figure 6, ‘Safe’ correlation with ‘Robot peak vel.’ $|\rho| < 0.15$). This may indicate differences in the interpretation of the robot’s ‘intentions’ when the movement velocity was high but late, e.g. “robot is cooperative and trying to catch up with me”, versus when the velocity was high and early, e.g. “the robot is short-tempered, in a bad mood”.

The ‘spatial performance’ related measures were generally less correlated with the qualitative ratings than the ‘timing performance’ related measures, suggesting that the spatial endpoint error was generally within the acceptable margin of accuracy that people find acceptable.

The results lead us to conclude that the speed of interaction may carry a greater weight for the subjective experience than the accuracy of reaching the human desired object exchange location. The degree to which this applies depends, however, on the freedom of movement of the human participant, as seen by the reduced correlations $|\rho| < 0.25$ in the ‘Lying under the ear’ configuration (Figure 6, bottom section).

4.2 Implications

Generally speaking, the design of a service robot interacting with humans, for the specific case of object exchange, should be focused more on speed control rather than on the high reaching precision. It might be interesting to see up to which level the precision criteria is less important than speed (what precision in cm or in degrees is necessary to maintain a sufficient perceived quality for a given velocity)?

In any case, such analysis would not permit to extract universal constraints since we have seen that the satisfaction also depends on the freedom of movement of the human participant. In an open space, it might be possible to get both interaction partners conveniently placed. However, the interaction area, and in particular in industrial settings such as the one covered here, may affect the freedom of movement of the participant. This would tend to require a precision and or speed adaptation according to the liberty of motion available for the human.

The robot kinematics profile tended to be less critical to the participant’s sense of safety than the exchange timing with the human. Once more, this turns to implicate a higher importance to the design of a robotic system able to react very quickly to the human actions (in particular adjusting the motion start and the time of reaching the exchange site location). The lower importance of the robot kinematics parameter may also afford more freedom in the design of the robot motion profile, as long as the timing previously mentioned is respected. Indeed the human may prefer to have a robot well synchronized with their action rather than reproducing a human like trajectory.

An interesting extension of the current study would be to study how the humans adjust their alignment with the robot behaviour (and inversely) through repeated interaction to map how users adapt to the characteristics of a service robot. This would require larger numbers of trials (during several days) to see how, by accumulating exchange experiences, both partners get more used to and better synchronized with the other (if the robot control strategy permits it).

Acknowledgments: This work was supported by the CogLaboration project within the 7th Framework Programme of the European Union, – Cognitive Systems and Robotics, contract Number FP7-287888, see also http://www.coglaboration.eu/.

We would like to thank the volunteers from Tecnalia who sacrificed their precious time to participate as subjects for our Human-Robot Interaction experiments.

REFERENCES

