

# Adaptive Coordination Strategies for Human-Robot Handovers

Chien-Ming Huang

Department of Computer Sciences  
University of Wisconsin–Madison  
cmhuang@cs.wisc.edu

Maya Cakmak

Department of Computer Science & Engineering  
University of Washington  
mcakmak@cs.washington.edu

Bilge Mutlu

Department of Computer Sciences  
University of Wisconsin–Madison  
bilge@cs.wisc.edu

**Abstract**—Handovers of objects are critical interactions that frequently arise in physical collaborations. In such interactions, humans naturally monitor the pace and workload of their partners and adapt their handovers accordingly. In this paper, we investigate how robots designed to engage in physical collaborations may achieve similar adaptivity in performing handovers. To that end, we collected and analyzed data from human dyads performing a common household task—unloading a dish rack—where receivers had different levels of task demands. We identified two coordination strategies that enabled givers to adapt to receivers’ task demands. We then formulated and implemented these strategies on a robotic manipulator. The implemented autonomous system was evaluated in a human-robot interaction study against two baselines that use “proactive” and “reactive” coordination methods. The results show a tradeoff between team performance and user experience when human receivers had greater task demands. In particular, the proactive method provided the greatest levels of team performance but offered the poorest user experience compared to the reactive and adaptive methods. The reactive method, while improving user experience over the proactive method, resulted in the poorest team performance. Our adaptive method maintained this improved user experience while offering an improved team performance compared to the reactive method. Our findings offer insights into the tradeoffs involved in the use of these methods and inform the future design of handover interactions for robots.

## I. INTRODUCTION

Many robotic technologies that are designed for domestic, workplace, and clinical use will engage in physical activities with people. Examples include a household robot that helps its user unload groceries, a collaborative manufacturing robot that hands a human worker parts for assembly, and a rehabilitation robot that guides patients through physical exercises. Such physical activities can be characterized as *joint actions* in which parties coordinate their actions in space and time to achieve a common goal [25]. To achieve seamless coordination, parties monitor each other’s task actions and progress and adapt their actions based on these observations [26]. Such awareness and adaptivity are critical to team performance [26] and the psychological consequences of the interaction [20].

*Handovers*, which involve the transfer of an object from a giver to a receiver, are fundamental joint actions that enable physical collaborations. Prior research has investigated how various aspects of the robot’s presentation of objects, including its use of differentiable motions [7], object affordances [9], appropriate approach angles [34], and social cues [21], con-

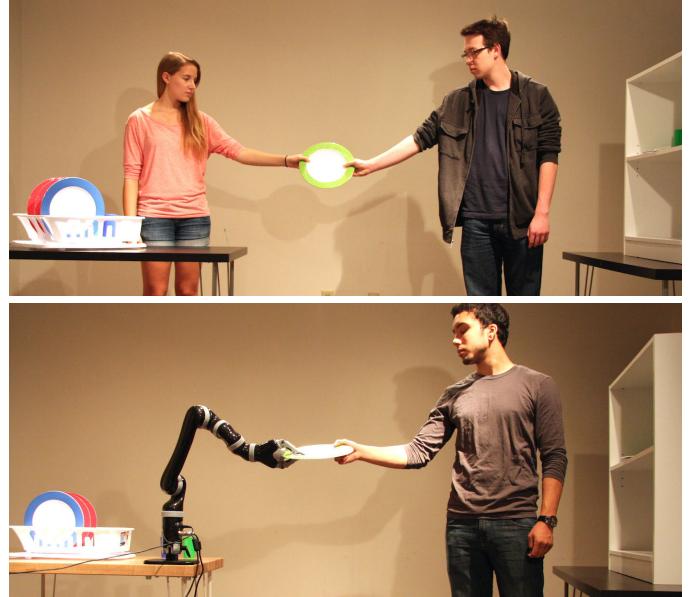


Fig. 1: We studied human-human handovers (top) in a household scenario, identified strategies that humans used for coordination, implemented them on a robotic manipulator, and evaluated their effectiveness in supporting coordination in human-robot handovers (bottom).

tribute to the success of human-robot handovers. How they may adapt to user task demands to achieve temporally aligned handovers remains unexplored. An awareness of user states, particularly of the user’s current focus, progress in the task, and availability, can enable robots to adaptively perform handovers, better supporting team performance and user experience.

In this paper, we seek to better understand adaptation strategies that enable humans to seamlessly coordinate their actions and to explore how robots may leverage such strategies to more successfully engage in physical interactions with their users. We contextualized our investigation in a household application scenario in which parties collaboratively unloaded dishes from a drying rack (Figure 1, Top). We collected data from pairs of human participants as they performed handover actions under different task demands. The analysis of this data resulted in a computational model of adaptive coordination that was implemented on a robotic manipulator. We conducted a human-robot interaction study in a similar collaborative scenario (Figure 1, Bottom) to evaluate the effectiveness of the model in supporting human-robot handovers.

In the next section, we review prior work on joint action and handovers in human-human and human-robot interactions. This review is followed by a description of our model of coordination in human-human handovers, the implementation of this model on an robotic manipulator, and the evaluation of the different coordination methods in a human-robot interaction study. We then discuss our findings and their implications for future work, concluding with a summary of our contributions.

## II. BACKGROUND

### A. Joint Action and Action Coordination

Physical collaborations require humans to engage in joint actions with their partners toward a common goal [26]. Advances in the understanding of joint action in human interactions have informed the development of supporting mechanisms for human-robot joint action (e.g., [13, 18, 22, 28]). For example, Hoffman and Breazeal [13] developed a computational model that generated anticipatory actions for an assistive agent to enable the agent to adapt to its user’s workflow in a simulated assembly scenario. They showed that an agent using anticipatory actions, compared to a reactive agent, formed a more fluid team with its users, resulting in greater concurrent activity with them. Inspired by coordination behaviors of human teams, Shah et al. [28] developed a plan execution system called *Chaski* that adapts to human partners and seeks to minimize its partners’ idle time. Cross training has also been explored as a strategy for designing effective human-robot teams [23]. These studies highlight the promise of drawing on human coordination strategies to achieve fluid human-robot teams.

While prior work highlights the implications that successful joint action has for improving team performance, user experience is another factor that designers must take into account when developing robots for physical collaboration. For instance, in domestic tasks such as doing daily chores such as unloading groceries, users may want to interact with robots at their own pace, as opposed to aiming to maximize team efficiency. Therefore, we postulate that solely maximizing task performance may not necessarily result in desirable joint action. We also highlight the importance of incorporating an awareness of the partner’s task into planning one’s own actions.

### B. Designing Handovers for HRI

*Handovers* involve the transfer of objects from a giver to a receiver and serve as a fundamental skill for complex collaboration and interaction. Prior work has investigated various aspects of *human-human handovers*, including velocity, force, and style, in order to inform the design of human-robot handovers. For instance, Becchio et al. [5] found that people display different velocity profiles when they hand over an object compared to when they place the object on a surface. Their characterization of handover motion profiles was consistent with earlier observations, which showed that people display minimum-jerk motions during handovers [14, 30]. Other work modeled grip force patterns during handovers to design handover controllers [8]. To design a robot that more effectively delivers flyers, Shi et al. [29] studied different styles

of handovers that shopkeepers used when distributing flyers to passersby in a mall. While these studies provide useful insights into the different facets of handovers, how people coordinate handover actions in situations where the receiver is distracted or delayed by a secondary task—a common occurrence in everyday situations—is unknown.

*Human-robot handovers* offer a rich design space with a large number of parameters. Research to date has explored the role of gaze [21], approach angle and saliency [34], contrast between start and end points of handover motion [7], and anthropomorphism [31]. Moreover, prior work developed methods to choose handover parameters such as pose and trajectory that consider user preferences [6], user comfort [2], object affordances [3, 9], and user mobility constraints [19] that facilitate human-robot handovers. Although most work focused on a robot handing objects to humans, how robots may take objects from humans has also been investigated [2, 10]. Previous research has also explored how robots may utilize handovers across a number of application scenarios, including manufacturing [15, 34], household service [10, 19], or serving drinks [7, 11]. Our work contributes to the exploration of the rich design space for handovers by investigating the effects of task demand on handover actions and developing new methods to enable robots to adapt to their users’ changing task demands.

## III. COORDINATION IN HUMAN-HUMAN HANDOVERS

In this section, we describe our study of human-human handovers that informed our design of coordination strategies to facilitate human-robot handovers.

### A. Data Collection

*Task & Setup* – To better understand how people adapt handover actions to task demands, we observed pairs of human participants collaboratively unload a dish rack (Figure 1, top). The task required one participant, the *giver*, to pick up plates and cups in a drying rack and hand them to the other participant, the *receiver*, who then placed these items on a nearby shelf. Participants performed two variations of this task. In the first variation, both the giver and receiver were engaged only in unloading dishes, denoted by *regular unloading* hereafter. In the second variation, in addition to unloading dishes, the receiver was given a secondary task of matching patterns on the items to specified target locations, denoted by *tasked unloading* hereafter. We introduced the secondary task in order to add to the receiver’s cognitive load [33], thereby increasing task demand, such that the giver would need to adapt to the receiver’s workload. Each dyad performed both task variations once. The roles of giver and receiver within a dyad were maintained throughout the interaction. Positions of participants’ body joints were recorded using Microsoft Kinect version 2. The interactions were also video-recorded. Participants received \$5 USD for their participation in the study.

*Participants* – Eight dyads, two for each gender combination, were recruited. Participant ages ranged 20–34 ( $M = 23.94$ ,  $SD = 4.58$ ). All dyads but one included participants who did not know each other prior to this study.

## B. Data Processing

Overall, we observed that givers (1) monitored the receivers' task progress, especially in the *tasked unloading* task, and (2) adapted to the receiver's pace by pausing and/or slowing down their actions. These observations informed our analysis for further understanding coordination strategies in handovers.

### 1) Annotating User State

We categorized the handover activity into six different states for the giver and the receiver. The giver's states include (1) *reach*: moving hand from a resting position or the center of body to grasp an object; (2) *retrieve*: grasping and moving the object to the center of body; (3) *give*: moving the object from the center of body to the handover position; (4) *handover*: both the giver and receiver touching the object and the giver releasing the object; (5) *retract*: moving the hand back to the center of body; and (6) *idle*: all other actions. Similarly, the receiver's states include (1) *take*: moving hand from the resting position or the center of body to the handover position; (2) *handover*: both the giver and receiver touching the object and the giver releasing the object; (3) *retrieve*: the giver releasing the object and the receiver moving it to the center of body; (4) *place*: moving the hand from the body center to hand off of the object; (5) *retract*: returning hand to the body center or a resting position; and (6) *idle*: all other actions.

The video data was annotated with these states for a detailed analysis of handover actions and for training an algorithm for online prediction of user states during interaction with a robotic manipulator (see Section IV-A). The data included 8389 and 12793 joint readings collected in the regular and tasked unloading conditions, respectively. A primary rater coded all of the data, and a secondary rater coded 10% of the data. Inter-rater reliability analysis showed substantial agreement between the raters (Cohen's  $\kappa = .74$ ) [16].

### 2) Smoothing Sensor Data

We applied an Exponentially Weighted Moving Average (EWMA), a common noise reduction technique for time-series data defined in Equation 1, to reduce noise in the raw joint position data from the Kinect sensor.

$$\mathbf{x}_t = \alpha \mathbf{y}_{t-1} + (1 - \alpha) \mathbf{x}_{t-1} \quad \text{for } t > 0, \mathbf{x}_0 = \mathbf{y}_0 \quad (1)$$

where  $\mathbf{y}_t$  is the raw sensor measurement of the joint positions in Cartesian space at time  $t$ , and  $\mathbf{x}_t$  is the filtered measurement at time  $t$ . The weighting parameter  $\alpha$  controls how much we discount prior data, which our testing determined to be 0.2 for best performance. The results reported below were based on analyses using smoothed data.

## C. Coordination Strategies

To understand how people coordinate their actions, we first need to know *when* coordination strategies are needed. An inspection of the average durations of receiver actions in each state in the regular and tasked conditions, as illustrated in Figure 2, showed that receivers stayed in the *retrieve* and *place* states longer in the tasked condition than in the regular condition. These differences suggest that these states were likely to be when givers had to adapt to the availability of the receivers.

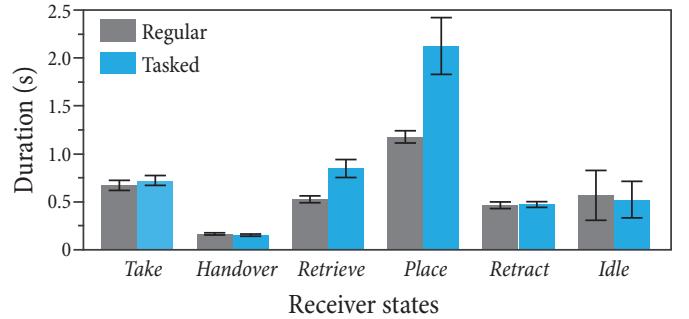


Fig. 2: Average durations of the receiver's states in the regular and tasked conditions. Receivers took longer in the *retrieve* and *place* states, suggesting that these were the states in which the giver had to adapt their actions to the receiver's availability. Error bars indicate 95% confidence intervals.

Our analysis next focused on giver behaviors in order to better understand the adaptations they displayed (Figure 3), identifying two adaptive strategies: *waiting* and *slowing down*.

### 1) Waiting strategy

In the tasked condition, the giver adapted his/her actions by *waiting* for the receiver to complete the secondary task before resuming unloading dishes. The giver was found in this idle state a total of 79 times (across 96 tasked trials). Waiting was most commonly observed, a total of 35 times, after the giver retrieved the item but before passing it to the receiver (e.g., keeping the item in front of body and ready for handover). The giver also waited after retracting from the handover pose but before reaching for the next object (24 times), and after reaching toward but before retrieving the object (17 times). In rare occasions, the giver waited in the handover pose (3 times). This waiting strategy is consistent with prior work that reported the giver pausing his/her action until the receiver was ready for handover [17].



Fig. 3: Velocity profiles and example snapshots from adaptive coordination strategies—slowing down (top) and waiting (bottom)—displayed by givers in human-human handovers.

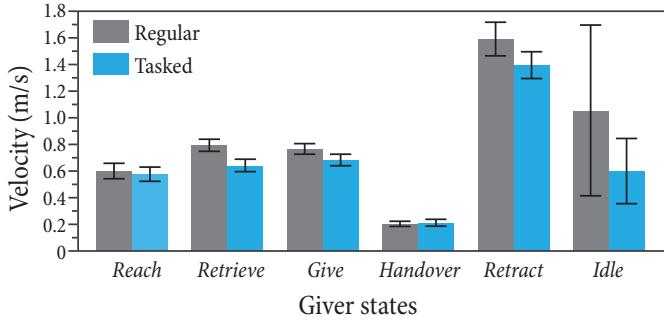


Fig. 4: Average velocities of the giver’s hand across different states in the regular and tasked conditions. Givers slowed down their actions during *retrieve*, *give*, and *retract* states. Error bars indicate 95% confidence intervals.

## 2) Slowing-down strategy

In addition to the waiting strategy, we observed that the giver adapted to the receiver’s availability by *slowing down* his/her actions while the receiver was occupied by the secondary task. We further inspected the velocity of the giver’s hand across the two task conditions (Figure 4) and found it to be slower during the *retrieve*, *give*, and *retract* states in the tasked condition than in the regular condition. These observations confirmed that the giver slowed down actions to adapt to the receiver’s availability and highlighted the states in which the giver slowed down.

We note that people used these two strategies in a combined and interchangeable way, as shown in Figure 3, and that slowing down usually occurred prior to waiting, indicating that givers slowed down their action and then paused if necessary.

## IV. ENABLING ADAPTIVE HUMAN-ROBOT HANDOVERS

In this section, we describe our design of coordination strategies to enable adaptive human-robot handovers based on the strategies we identified in the human-human interaction study. Our goal is to develop a human-aware controller that adapts to the availability of the user in planning robot actions for handover. To that end, we developed an autonomous system that uses human joint positions to predict the state that the user is currently in (Section IV-A) and determines what action the robot should take accordingly (Section IV-B).

### A. Predicting User State

Coordination in joint action requires knowledge of the ongoing actions and the current states of interaction partners [26]. To inform the robot of user actions and states, we employed a Microsoft Kinect version 2 camera to track the user’s body joints, extracted features from the body joints that represented characteristics of the user’s current action, and predicted the user’s current state using a K-Nearest Neighbor (KNN) algorithm. We provide details of this process below.

#### 1) Temporal and Spatial Features

Prior to extracting features, we applied an EWMA filter (Equation 1) to the raw joint data in order to smooth out sensor noise. Using the filtered data, we derived the following features to represent the spatial and temporal state of the user.

- *Hand velocity ( $\Delta x_t^{hand}$ )*: This feature captures the velocity of the user’s active hand. Note that the user might use

different hands in different states. For instance, the receiver may use the right hand to take the object and the left hand to place it. This feature was calculated using equation:

$$\Delta x_t^{hand} = \frac{x_t^{hand} - x_{t-1}^{hand}}{\Delta t} \quad (2)$$

where  $x_t^{hand} \in \mathbb{R}^3$  is the Cartesian position of hand\_left or hand\_right joints in the Kinect joint structure, and  $\Delta t$  is the duration between sensor readings (about 30ms).

- *Extension ( $x_t^{ext}$ )*: This feature represents the extension of the user’s arm as a vector from the origin of the body, denoted as  $x_t^{origin}$ , to the active hand.

$$x_t^{ext} = x_t^{hand} - x_t^{origin} \quad (3)$$

The position of the spine\_mid joint is used as  $x_t^{origin}$ .

- *Approach ( $x_t^{app}$ )*: This feature characterizes the extent to which the user’s arm approaches the other agent as a vector between the active hand and the midpoint of the body centers of the two parties. The basis of this feature was the observation that handovers happen approximately at the midpoint between two parties [4].

$$x_t^{app} = x_t^{hand} - x_t^{midpoint} \quad (4)$$

$$x_t^{midpoint} = \frac{x_t^{origin} + x_t^{origin\_other}}{2} \quad (5)$$

The state of the user is represented with a feature vector consisting of these three features,  $f = (\Delta x_t^{hand}, x_t^{ext}, x_t^{app})$ .

#### 2) State Prediction

We used a KNN classifier to predict the user’s current state based on the features described above. When a new observation arrives in the form of a feature vector, the algorithm finds the  $K$  most similar instances in the training dataset (the annotated data from the human-human interaction study) according to the distance measure in Equation 6.

$$d(f, \bar{f}) = \|\Delta x_t^{hand} - \bar{\Delta x}_t^{hand}\| + \|x_t^{ext} - \bar{x}_t^{ext}\| + \|x_t^{app} - \bar{x}_t^{app}\| \quad (6)$$

where  $f$  denotes the new observation to be classified;  $\bar{f}$  denotes a sample in the training dataset; and the distance between individual feature pairs are the  $L_2$  norms of the difference vector. The weights of the features in  $f$  were chosen to be equal, indicating that each feature contributes to the prediction equally, based on our preliminary testing. We set  $K$  to be 35 for this application (See Section IV-A3), i.e., the 35 most similar instances in the annotated data formed a group of candidate predictions. The prediction of the user’s state was based on the majority vote of the 35 candidates. The algorithm’s confidence for each prediction was calculated using Equation 7.

$$\text{Confidence} = \frac{\text{number of majority classifications}}{K} \quad (7)$$

We calculated average confidence scores for correct and incorrect predictions, shown in Table I, and used these scores as thresholds to filter out predictions with low confidence.

TABLE I: Confidence scores for KNN-based prediction of receiver states.

		Take	Handover	Retrieve	Place	Retract	Idle	Dropping %
Regular	Correct	92.03	68.07	89.77	94.39	90.27	83.99	38.44%
	Incorrect	72.96	64.64	70.11	75.52	71.13	75.94	17.80%
Tasked	Correct	89.93	61.74	88.02	92.46	86.84	92.95	40.64%
	Incorrect	68.29	56.25	65.29	74.16	69.22	62.97	16.62%

### 3) Parameter Selection and Evaluation

The parameters, including  $\alpha$  for the EWMA filter and  $K$  for the KNN algorithm, were tuned based on eight-fold leave-one-out cross-validation, using seven dyads for training and the remaining dyad for testing. Since the two datasets—*regular* and *tasked* unloading—yielded similar results (0.4% difference in accuracy), we chose to use the best parameters learned from the regular dataset. To assess the effectiveness of our KNN model in predicting the receiver’s state, we conducted another eight-fold leave-one-out cross-validation using the tasked dataset with the chosen parameters (Figure 5). The results of this test showed that our KNN model more accurately predicted receiver states than two baselines—*chance* and *most common guess* [1]—did. The *chance* baseline involves random guesses of user state, and the *most common guess* baseline always predicts the state that most commonly appeared in the training data.

In addition to the ordinary use of KNN, we explored confidence thresholding using the thresholds that we calculated for incorrect predictions. This method resulted in a tradeoff between improved accuracy and the number of dropped predictions; our results showed that confidence thresholding improved accuracy by approximately 6%, although approximately 17% of the predictions were dropped. The dropped predictions usually occurred during transitions between states. For the human-robot interaction study, reported in Section V, the algorithm dropped 15.35% of the predictions. The average duration of dropped predictions was 198.68 ms, while the algorithm made a prediction every 46.4 ms. We found the rate and duration of dropped predictions to be acceptable in our application, as our algorithm provided predictions at a high frequency.

### B. Generating Robot Actions

The coordination strategies that we observed in human-human interactions informed the development of a model to emulate human-style adaptive coordination, as outlined in Algorithm 1. Here, *userState* is provided by the KNN

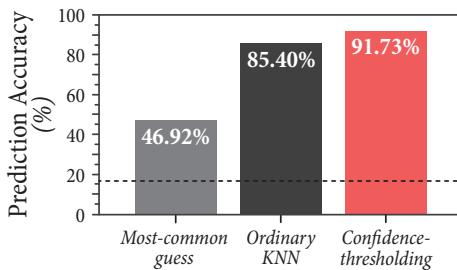


Fig. 5: Cross-validation results of our KNN model in predicting receiver states using the *tasked* dataset. The dashed line indicates baseline accuracy (16.67%).

model described above. The *delayThreshold* parameter was set to 1.36 seconds, which was the average duration for people to finish retrieving and placing an object when no secondary task was present. In Line 1 of Algorithm 1, the robot’s waiting position is determined based on a probability distribution obtained from our data on human-human interactions (Section III-C1). In Line 3, *currentRobotState* is specified by a deterministic finite state machine that represents the sequence of giver states. Our model was implemented in ROS [24] to control a Kinova MICO robotic arm, shown in Figure 1. We used force-sensor information from the robot’s joints to determine whether or not items in the robot’s gripper were grasped by the user and to plan for their release.

## V. EVALUATION OF HUMAN-ROBOT HANDOVERS

In this section, we report on a study of the effects of adaptive coordination on objective and subjective outcomes of human-robot handovers. In particular, we evaluated the effectiveness of the coordination strategies described above in improving task performance and user perceptions in handover interactions where a robotic manipulator passed objects to human participants in a common household scenario.

### A. Hypothesis

Our evaluation tested the central hypothesis stated below regarding the effects of a robot’s use of human-inspired coordination strategies on team performance and perceptions of the robot under different levels of task demand.

*Hypothesis* — When users are under high levels of task demand, the robot employing coordination strategies that enable it to adapt its handover actions to its user’s task will improve team performance and user experience with and perceptions of the robot, while employing these strategies will not offer similar benefits when users are under low levels of task demand.

### B. Experimental Design, Task, & Conditions

To test our hypothesis, we conducted a  $3 \times 2$  within-participants study in which we manipulated the *coordination method* that the robot employed and the level of *task demand* under which the participants worked. The paragraphs below describe the three coordination methods considered in the study.

---

#### Algorithm 1 Adaptive Coordination Strategy for Handover

---

**Require:** *userState*, *delayThreshold*

- 1: probabilistically select a robot waiting position
- 2: **while** ISCURRENTHANDOVERTRIALACTIVE( ) **do**
- 3:     *currentRobotState*  $\leftarrow$  GETROBOTSTATE( )
- 4:     *robotAction*  $\leftarrow$  GETROBOTACTION(*currentRobotState*)
- 5:     **if** *userState* = *retrieve* or *place* **then**
- 6:         **if** ELAPSEDSTATETIME( )  $\geq$  *delayThreshold* **then**
- 7:             **if** ISROBOTATWAITINGPOSITION( ) **then**
- 8:                 WAIT( )
- 9:             **else**
- 10:                 SLOWEXECUTION(*robotAction*)
- 11:             **end if**
- 12:         **end if**
- 13:     **end if**
- 14:     REGULAREXECUTION(*robotAction*)
- 15: **end while**

---

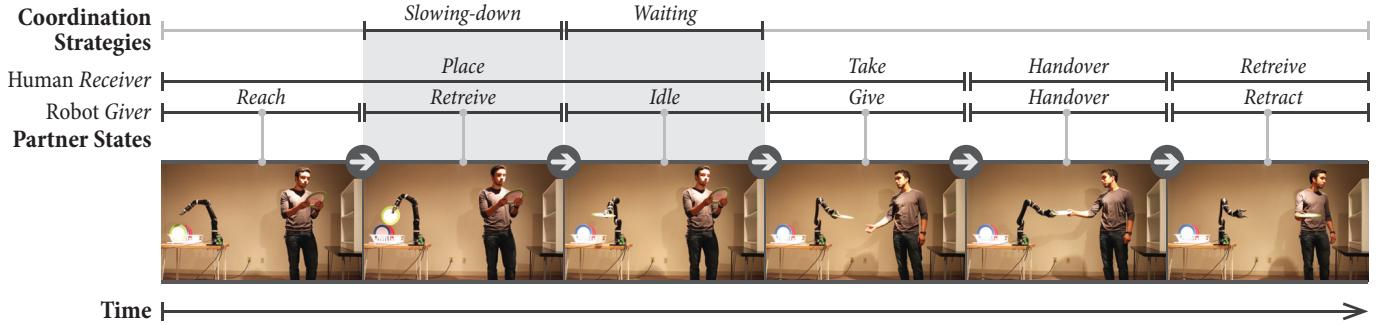


Fig. 6: An example human-robot handover using our method for adaptive coordination that employed the *waiting* and *slowing-down* strategies.

**Proactive coordination** — Following this method of coordination, the robot did not take the user’s task demand into account when planning its actions. After handing an object to the user, the robot *proactively* fetched the next object and presented it to the user, even if the user was not ready to take the next object. Using this method, the robot aimed to minimize the user’s idle time and to maximize concurrent activity with the user, measures proposed by prior work as objective indicators of fluid teamwork [12, 23, 27, 28].

**Reactive coordination** — In contrast to the proactive coordination, the robot following this method waited for its user to fully complete the task, thus *reacting* to the user’s availability. When the user returned to an idle state, the robot fetched the next object. This method enabled a turn-based interaction in which interaction partners asynchronously contributed to the task. Prior work has characterized this method of coordination as a form of joint action [22].

**Adaptive coordination** — The robot achieved adaptive coordination by following Algorithm 1, which enabled the robot to utilize the *waiting* and *slowing-down* strategies to adapt to its user’s task demand.

Moreover, we introduced a manipulation to the experimental task to create two levels of task demand as described below.

**Low task demand** — In this setting, participants were asked to engage in the task of unloading dishes, where the robot passed four plates and two cups to its user, and the user placed the objects on a shelf, as shown in the bottom of Figure 1. Because people can take and place an object faster than it takes the robot to fetch and deliver the object, this setting creates a lower level of task demand for the user.

**High task demand** — In this setting, in addition to the task of unloading dishes, participants were engaged in a secondary task, which involved solving a math problem attached to each object and placing the object in a location on the shelf that corresponds to the answer. Each problem consisted of nine single-digit numbers and involved all four arithmetic operations (i.e., addition, subtraction, multiplication, and division) twice (e.g.,  $7 \times 2 - 2 + 6 \div 3 - 9 + 2 \times 8 \div 4$ ). These problems were designed to ensure that each problem had a similar level of difficulty. The same set of 18 unique math problems was used for each participant. The problems were randomly assigned to objects and to coordination methods in order to prevent any systematic bias due to differences in problem difficulty. In each

round of interaction, nine potential answers, six of which were correct for the six handover actions, were provided along with the option “not able to find the answer.” The design of this task aimed to create a higher level of task demand for the user.

The three coordination methods and two task-demand levels comprised six experimental conditions. The robot followed the same pre-programmed motion trajectories across all conditions.

### C. Measures

We used objective and subjective measures to assess the effectiveness of the three coordination methods. Objectively, we measured *team performance* as informed by prior work [12, 23, 28]. In particular, we measured task completion time (the time between the robot’s first move to pick up the first object and the user placing the last object), concurrent activity (the proportion of the total time that both the user and the robot were in action to the total task completion time), user idle time (the proportion of task completion time to total time of inaction by the user or the robot), and robot idle time.

In addition to objective measures, we developed four scales—*fluency*, *intelligence*, *awareness*, and *patience*—to measure participants’ experience with and perceptions of the robot. The fluency scale extended a previously proposed measure of subjective fluency [12] and consisted of five items (Cronbach’s  $\alpha = 0.91$ ). The scales of intelligence, awareness, and patience consisted of four (Cronbach’s  $\alpha = 0.88$ ), two (Cronbach’s  $\alpha = 0.83$ ), and four (Cronbach’s  $\alpha = 0.85$ ) items, respectively. All items were on a seven-point rating scale (1 = Strongly Disagree, 7 = Strongly Agree).

### D. Procedure

Following informed consent, participants spent a minute to review the order of operations and the multiplication table prior to beginning the task. The study involved six rounds of interaction—one round for each condition. The order of coordination methods and levels of task demand were counterbalanced. After each round of interaction, participants filled out a questionnaire regarding their experience with and perceptions of the robot. Finally, the experimenter conducted a post-experiment interview. The study took approximately 50 minutes. All participants received \$10 USD as compensation.

### E. Participants

A total of 26 participants were recruited from the local community. However, two participants were excluded from the

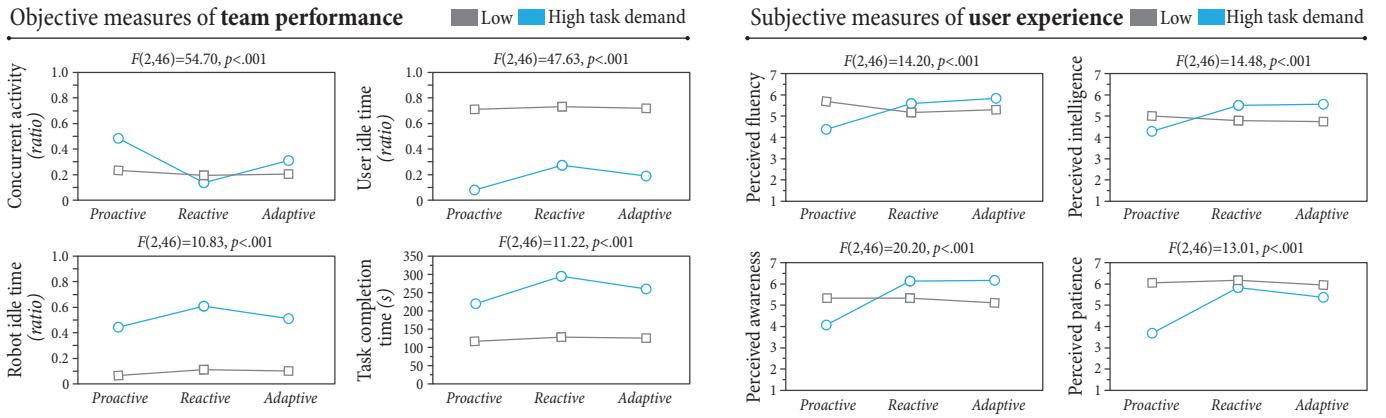


Fig. 7: Interaction plots and ANOVA test details for objective measures of team performance (left) and subjective measures of user experience (right).

data analysis, as one of the them did not finish the task, and the other one failed to follow instructions. The resulting 24 participants (13 females, 10 males, and one unspecified) were aged between 18 and 35 years ( $M = 20.54, SD = 3.72$ ) and reported little familiarity with robots ( $M = 2.79, SD = 1.47$  on a seven-point scale). None of the participants in this study had taken part in the human-human data collection study.

#### F. Results

Our analysis of the data from each measure involved a two-way repeated-measures analysis of variance (ANOVA), using coordination method, task demand, and their interaction as independent variables and participant ID as a random variable. Guided by our hypothesis, our analyses focused the interaction between coordination method and task demand. Six *a priori* pairwise comparisons, using a Bonferroni-adjusted  $\alpha$  level of .008 (.05/6) for significance, were carried out to identify differences across conditions for each objective and subjective measure. For readability, details of statistical tests are omitted from the text, and test details for interaction effects and pairwise comparisons are provided in Figures 7 and 8, respectively.

##### 1) Objective Measures

Across all objective measures, our analyses revealed significant interaction effects between coordination method and task demand. At lower levels of task demand, pairwise comparisons showed no differences in any measure of team performance across coordination methods. In contrast, at high levels of task demand, the coordination method that the robot used had a significant effect on the task performance of the human-robot team in measures of task completion time, concurrent activity, user idle time, and robot idle time. In particular, our results showed that, among the three methods, proactive coordination yielded the greatest outcomes in our task performance measures, followed by adaptive coordination, while reactive coordination resulted in the poorest outcomes.

##### 2) Subjective Measures

Similar to the objective measures, our analyses revealed significant interaction effects between coordination method and task demand for all subjective measures. Comparisons showed that, across different coordination methods, there were

no significant differences in participants' perceptions of team fluency and of the robot in terms of intelligence, awareness, and patience when they were under high levels of task demand. However, differences in subjective measures emerged when participants were under high levels of task demand during their interactions with the robot.

Interestingly, while objective outcomes of team performance in measures of concurrent activity, user idle time, and robot idle time indicated a better performance when the robot used the proactive coordination method, the subjective measure of fluency indicated otherwise. Participants perceived their interactions with the robot using the proactive coordination method to be less fluid compared to the other two coordination methods. Both reactive and adaptive coordination yielded a similar degree of perception of team fluency. Our analyses of the data from measures of perceived intelligence, awareness, and patience of the robot were consistent with these findings.

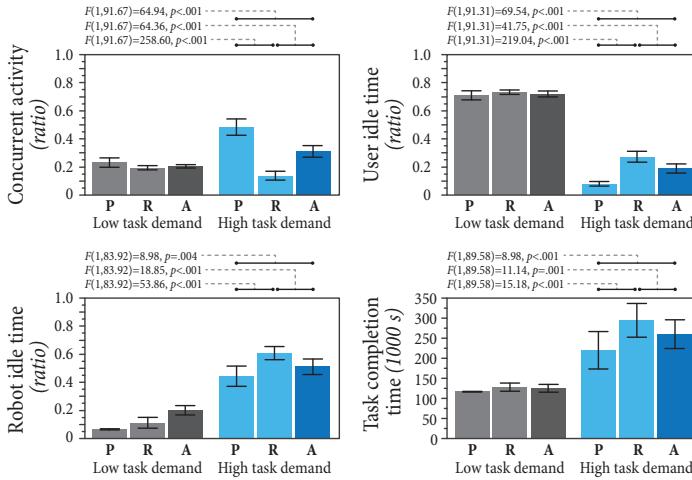
## VI. DISCUSSION

We identified two coordination strategies, *waiting* and *slowing down*, from data on human-human handovers in a household application scenario and developed an adaptive coordination method involving these two strategies. We implemented a robot system that autonomously performed handovers with users in a similar scenario and evaluated the effectiveness of our adaptive coordination method against two alternative methods in facilitating human-robot handovers. Our results showed a tradeoff between team performance and user experience. Below, we discuss our results and their implications for designing effective coordination strategies for human-robot collaboration.

### A. Design Implications

Our results showed that the effects of coordination methods were differentiable only when participants' task demand was greater than that of the robot, suggesting that a robotic assistant should monitor the task progress of its user and employ coordination methods as appropriate and when necessary. Our results further revealed a tradeoff between task performance and user experience based on the coordination method employed. While proactive coordination significantly improved team performance in measures of concurrent activity and idle time, it

## Objective measures of team performance



## Subjective measures of user experience

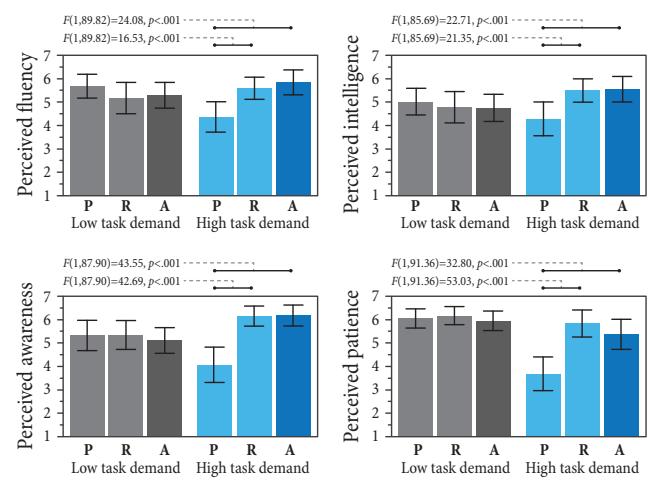


Fig. 8: Data from measures of team performance and user experience. P, R, and A represent the *proactive*, *reactive*, and *Adaptive* coordination methods, respectively. Pairwise comparisons use a Bonferroni-adjusted  $\alpha$  level of .008 for significance. Error bars indicate 95% confidence intervals.

impaired users' experience with and perceptions of the robot. In line with prior work, this result indicates that task efficiency and user perceptions of team fluency are not necessarily positively correlated [13]. It also suggests a lack of correlation between perceived and objective measures of team fluency, as suggested by prior work on human-robot collaboration [12, 28]. We therefore argue that team fluency is a complex construct that could not be defined only objectively but that it requires a consideration of partners' perceptions and their task demands.

Team experience and performance are equally essential elements of joint action [20]. We found that proactive coordination improved team performance but hurt user experience. We speculate that this result was due to the pressure that the proactive robot may have imposed on users to complete their secondary task, as suggested by the excerpt below.

*It [the robot] moved really quickly and I was holding one plate still trying to figure out the problem and feeling like I have to pick up the next plate, so that was stressful.*

Conversely, reactive coordination improved user experience but reduced team performance, as indicated below.

*It [the robot] waited until I put the plate down to move again, hmm and that was almost too slow because then I'd have finished the math problem and be waiting to get the next one.*

Finally, adaptive coordination improved team performance while maintaining the support that reactive coordination provided for user experience, as predicted by prior research on human-human joint action [20, 26].

### B. Limitations

This work has limitations that motivate future investigation. First, we focused on a particular handover application that involved interaction partners unloading dishes. While we speculate that the coordination strategies of waiting and slowing-down will be applicable to other applications, the parameters that we derived from the collected data may not be

applicable to new applications. This work, however, illustrates a process for building coordination mechanisms for interactive robots. Second, although we demonstrated the potential of a KNN algorithm in predicting user state, its effectiveness and efficiency depend on the size of the training dataset. Although a larger dataset promises better accuracy, it may decrease efficiency due to the number of comparisons that the algorithm has to perform on the fly. Future work can explore alternative models such as decision trees (e.g., [32]) and probabilistic graphical models (e.g., [11]) for effective, efficient prediction of user states. Finally, the robotic manipulator used in this work placed constraints on the different characteristics of the motions it produced, including trajectories, velocity, and noise.

## VII. CONCLUSION

In human physical collaborations, parties adaptively coordinate their actions in order to achieve more fluid interactions and greater team performance and experience. This work contributes to the development of new methods for enabling similarly fluid and effective human-robot handovers in three ways. First, it offers a computational understanding of how people adapt their handover actions to the workload of their partners. Second, it demonstrates an autonomous robotic manipulator system that takes into consideration a real-time awareness of the task status of its user in performing handover actions. Third, it shows that use of different coordination methods results in a tradeoff between team performance and user experience. Our work also offers insights into the rich and nuanced design space for mechanisms that facilitate joint action in human-robot teams.

## ACKNOWLEDGEMENTS

This work was supported by National Science Foundation grants 1149970, 1208632, and 1426824. We would like to thank Brandi Hefty, Ethan Jesse, Christopher Ward, Jing Jing, and Catherine Steffel for their contributions.

## REFERENCES

- [1] Henny Admoni and Brian Scassellati. Data-driven model of nonverbal behavior for socially assistive human-robot interactions. In *Proc. ICMI'14*, 2014.
- [2] Jacopo Aleotti, Vincenzo Micelli, and Stefano Caselli. Comfortable robot to human object hand-over. In *Proc. RO-MAN'12*, 2012.
- [3] Jacopo Aleotti, Vincenzo Micelli, and Stefano Caselli. An affordance sensitive system for robot to human object handover. *International Journal of Social Robotics*, 6: 653–666, 2014.
- [4] Patrizia Basili, Markus Huber, Thomas Brandt, Sandra Hirche, and Stefan Glasauer. Investigating human-human approach and hand-over. In *Human Centered Robot Systems*, pages 151–160. 2009.
- [5] Cristina Becchio, Luisa Sartori, and Umberto Castiello. Toward you the social side of actions. *Current Directions in Psychological Science*, 19(3):183–188, 2010.
- [6] Maya Cakmak, Siddhartha S Srinivasa, Min Kyung Lee, Jodi Forlizzi, and Sara Kiesler. Human preferences for robot-human hand-over configurations. In *Proc. IROS'11*, 2011.
- [7] Maya Cakmak, Siddhartha S Srinivasa, Min Kyung Lee, Sara Kiesler, and Jodi Forlizzi. Using spatial and temporal contrast for fluent robot-human hand-overs. In *Proc. HRI'11*, 2011.
- [8] Wesley P Chan, Chris AC Parker, HF Machiel Van der Loos, and Elizabeth A Croft. Grip forces and load forces in handovers: implications for designing human-robot handover controllers. In *Proc. HRI'12*, 2012.
- [9] Wesley P Chan, Yohei Kakiuchi, Kei Okada, and Masayuki Inaba. Determining proper grasp configurations for handovers through observation of object movement patterns and inter-object interactions during usage. In *Proc. IROS'14*, 2014.
- [10] Aaron Edsinger and Charles C Kemp. Human-robot interaction for cooperative manipulation: Handing objects to one another. In *Proc. RO-MAN'07*, 2007.
- [11] Elena Corina Grigore, Kerstin Eder, Anthony G Pipe, Chris Melhuish, and Ute Leonards. Joint action understanding improves robot-to-human object handover. In *Proc. IROS'13*, 2013.
- [12] Guy Hoffman. Evaluating fluency in human-robot collaboration. In *Proc. HRI'13 Workshop on Human Robot Collaboration*, 2013.
- [13] Guy Hoffman and Cynthia Breazeal. Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. In *Proc. HRI'07*, 2007.
- [14] Markus Huber, Markus Rickert, Alois Knoll, Thomas Brandt, and Stefan Glasauer. Human-robot interaction in handing-over tasks. In *Proc. RO-MAN'08*, 2008.
- [15] Ansgar Koene, Anthony Remazeilles, Miguel Prada, Ainara Garzo, Mildred Puerto, Satoshi Endo, and Alan M Wing. Relative importance of spatial and temporal precision for user satisfaction in human-robot object handover interactions. In *Proc. New Frontiers in Human-Robot Interaction '14*, 2014.
- [16] J Richard Landis and Gary G Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174, 1977.
- [17] Min Kyung Lee, Jodi Forlizzi, Sara Kiesler, Maya Cakmak, and Siddhartha Srinivasa. Predictability or adaptivity?: Designing robot handoffs modeled from trained dogs and people. In *Proc. HRI'11*, 2011.
- [18] Claus Lenz, Suraj Nair, Markus Rickert, Alois Knoll, W Rosel, Jürgen Gast, Alexander Bannat, and Frank Wallhoff. Joint-action for humans and industrial robots for assembly tasks. In *Proc. RO-MAN'08*, 2008.
- [19] Jim Mainprice, Mamoun Gharbi, Thierry Siméon, and Rachid Alami. Sharing effort in planning human-robot handover tasks. In *Proc. RO-MAN'12*, 2012.
- [20] Kerry L Marsh, Michael J Richardson, and RC Schmidt. Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, 1(2):320–339, 2009.
- [21] AJung Moon, Daniel M Troniak, Brian Gleeson, Matthew KXJ Pan, Minhua Zeng, Benjamin A Blumer, Karon MacLean, and Elizabeth A Croft. Meet me where i'm gazing: how shared attention gaze affects human-robot handover timing. In *Proc. HRI'14*, 2014.
- [22] Bilge Mutlu, Allison Terrell, and Chien-Ming Huang. Coordination mechanisms in human-robot collaboration. In *Proc. HRI'13 Workshop on Collaborative Manipulation*, 2013.
- [23] Stefanos Nikolaidis and Julie Shah. Human-robot cross-training: computational formulation, modeling and evaluation of a human team training strategy. In *Proc. HRI'13*, 2013.
- [24] Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y Ng. Ros: an open-source robot operating system. In *Proc. ICRA'09 Workshop on Open Source Software*, 2009.
- [25] Natalie Sebanz and Guenther Knoblich. Prediction in joint action: what, when, and where. *Topics in Cognitive Science*, 1(2):353–367, 2009.
- [26] Natalie Sebanz, Harold Bekkering, and Günther Knoblich. Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, 10(2):70–76, 2006.
- [27] Julie Shah and Cynthia Breazeal. An empirical analysis of team coordination behaviors and action planning with application to human-robot teaming. *Human Factors*, 52(2):234–245, 2010.
- [28] Julie Shah, James Wiken, Brian Williams, and Cynthia Breazeal. Improved human-robot team performance using chaski, a human-inspired plan execution system. In *Proc. HRI'11*, 2011.
- [29] Chao Shi, Masahiro Shiomi, Christian Smith, Takayuki Kanda, and Hiroshi Ishiguro. A model of distributional handing interaction for a mobile robot. In *Proc. RSS'13*, 2013.
- [30] Satoru Shibata, Benlamine Mohamed Sahbi, Kanya

- Tanaka, and Akira Shimizu. An analysis of the process of handing over an object and its application to robot motions. In *Proc. SMC'97*, 1997.
- [31] Prasanna Kumar Sivakumar, Chittaranjan S Srinivas, Andrey Kiselev, and Amy Loutfi. Robot-human handovers in non-anthropomorphic robots. In *Proc. HRI'13*, 2013.
  - [32] Kyle Strabala, Min Kyung Lee, Anca Dragan, Jodi Forlizzi, and Siddhartha S Srinivasa. Learning the communication of intent prior to physical collaboration. In *Proc. RO-MAN'12*, 2012.
  - [33] J Ridley Stroop. Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6):643, 1935.
  - [34] Vaibhav V Unhelkar, Ho Chit Siu, and Julie A Shah. Comparative performance of human and mobile robotic assistants in collaborative fetch-and-deliver tasks. In *Proc. HRI'14*, 2014.