Proceedings of the 27th IEEE International Symposium
on Robot and Human Interactive Communication,
Nanjing, China, August 27-31, 2018

Designing Multimodal Intent Communication Strategies for Conflict Avoidance in Industrial Human-Robot Teams

Miles C. Aubert, Hayden Bader and Kris Hauser

Abstract— **Robot-to-human intent communication has been proposed as a method of enabling fluent coordination in human-robot teams. Prior research has focused on identifying modalities by which intent information can be accurately communicated, but has not yet studied whether intent communication enables fluent or safer coordination in human-robot teams in which intent communication is only supportive to the team's primary task. To address this question, we conduct a study (N = 29) in a mock collaborative manufacturing scenario in which motion-based and display-based intent communication approaches are evaluated under varying penalties for failing to coordinate safely. Subjective and objective measures of team fluency suggest that although intent communication supports fluent coordination, using a purely motion-based or a purely display-based approach may not be the most effective strategy. Although multimodal intent communication did not significantly improve upon unimodal approaches, merging both motion-based and display-based intent communication seems to combine the strengths of both approaches. Interestingly, results also suggest that contrary to theoretical predictions, the positive effect of intent communication is generally robust to teaming scenarios that require members to operate concurrently.**

I. Introduction

The communication of information and intended plans are important for coordinating teams of agents. Both the manner in which agents communicate and the impact of that communication on coordination have been studied extensively in human-human [1], [2] and human-robot teams [3], [4]. In human-robot teams, human workers often struggle to select appropriate coordination strategies because robots often do not follow human-like patterns of decision-making or communication [5]. To cope with this, human team members will often exhibit counterproductive behaviors (e.g. overcautiousness or incautiousness), which can negatively affect the performance and safety of collaboration [6], [7].

Fluent human-robot coordination may be facilitated by methods that allow a robot to directly support human teammates in selecting complimentary coordination strategies [6]. Intent communication has been used to help humans understand the internal state and future actions of robots [8]–[15], which may assist in the selection of complementary collaborative actions. However, communication comes at a cost. It has been argued that the act and comprehension of communication require an individual to devote some of their limited cognitive and temporal resources away from their primary task [16], [17]. If the human becomes distracted or unmotivated in attending to or acting upon the robot's communications, coordination may be negatively affected. Hence, it is important to design robot communication strategies structures to achieve adequate safety and performance.

M. C. Aubert, Hayden Bader and Kris Hauser are with the Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708 USA (phone: 919-619-5078; e-mail: miles.aubert@duke.edu).

In this paper we present an experimental investigation of intent communication within human-robot teams to address two open questions: (1) in teaming scenarios where human members must operate concurrently, dividing their resources between task-specific and coordination-specific activities, can intent communication persistently facilitate fluent coordination? and (2) if so, which modality best communicates intent information when human team members are switching between these two activities?

Our experiment evaluates two intent communication modalities (motion and display) within a dyadic human-robot team that asks participants to collaboratively complete a mock manufacturing task. The study (N = 29) evaluates two modalities of intent communication: (1) *motion*, whereby intent is communicated through the motion of the robot, and (2) *display*, whereby intent is explicitly communicated through an auditory and visual display. In this paper we refer to "modality" as the means of communication (motion or display), not the sensory channels over which each modality communicates intent. We evaluate these modalities both unimodally and multimodally, under varying levels of conflict penalties, which modulate the negative consequences for failing to coordinate. We do this to better understand the effects of intent communication under concurrent operations, where human team members may be less motivated to interpret and act upon the robot's intent. In high penalty scenarios, humans will be motivated to devote resources to coordination-specific activities, while in low penalty scenarios, humans will be less motivated to devote resources toward coordination, focusing their effort on task-specific activities. These techniques are evaluated on both objective and subjective measures of team fluency. Our results present statistically significant evidence that intent communication persistently facilitates fluent coordination in scenarios that require members to operate concurrently. Further, results show that in communicating intent a multimodal approach does not significantly outperform a purely motion-based or a purely display-based approach. However, there are indications that combining motion-based and display-based approaches may combine the strengths of both modalities.

II. Related Work

A. Intention and Coordination

The success of a team is often judged by its ability to adapt to rapidly changing task demands [1], [2], [18], [19]. To operate under evolving task demands, a team should select coordination strategies that allow team members to act collaboratively and minimize conflict [19]. To select

978-1-5386-7980-7/18/$31.00 ©2018 IEEE

coordination strategies that enable team members to fluently coordinate their actions with the actions of a teammate, some shared understanding of the teams behavior is required [1], [18], [20], [21]. It is widely accepted that the quality of this shared understanding is directly correlated with the fluency of coordinated action within a team [1], [18], [20], [21]. This shared understanding enables individuals to maintain an accurate prediction of their teammate's future actions and required resources, which in turn enables each team member to identify complementary actions that will minimize conflict and facilitate fluent coordination [1], [18], [22].

In achieving an accurate prediction of an individual's behavior, one must have some understanding of the underlying intent of that individual to exhibit a given behavior. The concept of intention is defined by the Theory of Reasoned Action [23] as "a person's subjective probability that he will perform some behavior." The theory of planned behavior [24] further suggests that if an individual is able to understand a teammate's intent, this will allow for the better prediction of their required resources and future actions, and therefore enable more fluent coordination.

In human-robot interaction there may be difficulties for a human to infer the intent of a robotic teammate due to inherent assumptions about the humanness and capabilities of the robot [4], [5]. When robot intent is difficult to infer, fluent coordination requires human team members to devote more resources to coordination-specific activities. In scenarios where humans are unable or unmotivated to devote sufficient resources toward coordination-specific activities they often adopt counterproductive interaction behaviors (e.g. being overcautious or being incautious), that can negatively impact the performance and safety of collaborative operations. Therefore, industrial designers ought to consider methods of augmenting a robot's ability to convey its intent, and in turn support the development of a shared understanding of the robot's internal state, capabilities, and future actions.

B. *Intent Communication*

To facilitate fluent coordination, researchers have suggested giving robots the ability to implicitly or explicitly communicate their intent to human teammates through modalities such as their motion [10], [11] or some form of display [12], [13]. This group of techniques is often referred to as intent communication [25]. The goal of robot-to-human intent communication is to express intent information through some modality that will allow a human teammate to more accurately predict the probability that their robotic teammate will perform a given behavior.

Motion-based communication can be subdivided into gesture [8], [9], [14], [26], [27] and legible motion [10], [11]. Traditional robot motion is "purely" functional [11], in which the motion is generated to achieve a specified goal without consideration of the teammate's understanding. Gestures are non-goal-oriented motions generated to explicitly communicate information to human teammates (e.g. pointing). Gestures can be used in both the planning and execution phases of a task to support humans in selecting complimentary collaborative actions. The concept of legible motion is defined as goal-orientated motion from which intent can be inferred. Legible motion can be organic where by a planned motion already communicates intent, such as when a car slows down ahead of a crosswalk when its intent is to wait for a crossing pedestrian. Legible motion can also be synthetic whereby motions are specifically generated to communicate intent by planning motions that exaggerate directionality early, which helps to disambiguate between goals. To date, existing work toward gestures and legible motion has experimentally shown the ability of these techniques in communicating intent [8]–[11], [14], [26], [27] but not in enabling fluent coordination. For example, the work by Dragan et al. in [11] presents an experimental evaluation of legible motion in a teaming scenario that asks participants to collaboratively perform a task by first interpreting the intent of their robotic teammate and then attending to the team task. Although their results highlight the utility of motion-based intent communication in enabling human team members to interpret a robotic teammate's intent, the format of their experiment does not consider scenarios in which team members must operate concurrently, switching between task-specific and coordination-specific activities. These scenarios may render humans unable or unmotivated to devote appropriate resources toward interpreting intent, and hence negatively impact the efficacy of intent communication in enabling fluent coordination.

Display-based intent communication directly conveys the intent of the robot through modalities such as visual [12], [15], auditory [13] or tactile [13]. Although these explicit displays can be diverse in their implementation [12], [13], [15], in general they aim to enable humans to understand when a robot will exhibit a given behavior and what that behavior will comprised of [12], [13], [15]. However, these techniques have been used with varying levels of success. For example, a study showed that humans have been able to determine the intent of a mobile robot with an expressive light display [15], but the use of an intent display did not aid humans in determining the intent of an autonomous vehicle [12]. The experimental evaluations of [12], [13], [15] were all formulated in a similar manner to the work in [11], such that, the experimental scenario was not necessarily an evaluation of the ability of intent communication to enable coordinated activities, but instead an evaluation of the ability of the proposed techniques to communicate intent accurately. For example, the work by Baraka et al. in [15] evaluates the use of expressive lights in communicating the intent of a mobile robot in an experiential scenario that asked participants to explicitly report the current state and actions of the robot from a predefined list of actions and states.

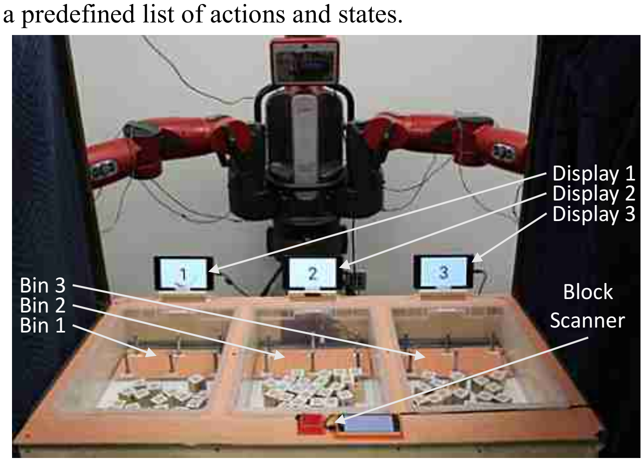
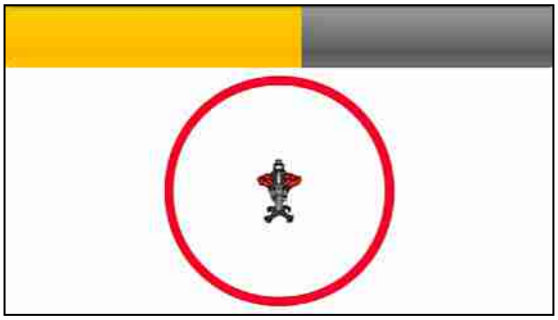


Figure 1. Photograph of simulated human-robot industrial workspace

(A)



(B)

Figure 2. Screenshots of the block scanner device. Assembly task details (A) and conflict notification (B).

More complex teaming scenarios such as collaborative manufacturing will require robotic team members to support their human teammates in selecting appropriate strategies that will allow them to coordinate fluently and safely in concurrent operations. Unlike prior work, humans will need to alternate between task-specific and coordination-specific activities. Coordination-specific activities, like withdrawing from a shared workspace or helping the robot complete its task, will require cognitive and temporal resources to be diverted from task-specific activities.

III. METHODOLOGY

We conducted a human-subject investigation (N = 29) of intent communication in a mock industrial scenario. Our experimental hypotheses are as follows:

Hypothesis 1 (H1): Intent communication (both unimodal and multimodal) will yield significantly better coordination in terms of both objective and subjective measures of team fluency in comparison to a control condition.

Hypothesis 2 (H2): Intent communication (both unimodal and multimodal) will have a smaller effect on both objective and subjective measures of team fluency when the negative consequences for failing to coordinate are less severe.

Hypothesis 3 (H3): Multimodal intent communication will yield a significant improvement in coordination over the two unimodal approaches in terms of both objective and subjective measures of team fluency.

The experiment is conducted as a within-subject design with two independent variables in a 4x2 structure for a total of 8 experimental trials. The two independent variables in our study are: (1) intent communication modality and (2) conflict time penalty.

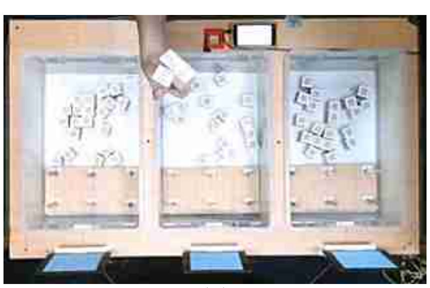
A. Simulated Collocated Industrial Human-Robot Team

We developed a simulated industrial workspace (Fig. 1) in which participants and a collocated Baxter robot perform a mock manufacturing task. The workspace consisted of:

1. Three assembly bins each containing 20 1.5-inch wooden cubes each labeled with a 3-digit number and 2D barcode.
2. An android tablet device positioned behind each of the three assembly bins facing the human participants used to show the visual intent displays.
3. A block scanner consisting of one Android phone device and a tabletop barcode scanner. The scanner is placed between the human and the assembly bins to provide task information, allow participants to interact with the robot, and answer post-trial surveys.

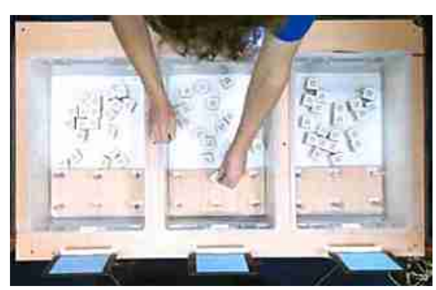
Participants are asked to perform twelve simple assembly cycles (Fig. 6) as quickly as possible while avoiding conflicts with the robot. A cycle consists of three phases as shown in Fig. 3. These phases are (1) Search, in which participants find a specified numbered block in a specified assembly bin as presented on the block scanning device as shown in Fig. 2A; (2) Scan, in which participants place the located block in the tabletop scanner to verify the correct block identity, and (3) Assemble, in which participants bolt the identified block onto an assembly peg specified on the block scanner device. The assembly phase requires unscrewing a nut, placing the block on the peg, and re-fastening the nut until snug.

Conflict in our task is defined as a situation in which both the robot and the human operate in the same assembly bin at the same time. The experiment is controlled so that conflict can occur in 50% of the cycles, forcing both the robot and the participant to operate in the same assembly bin. If a conflict occurs, the robot immediately stops and participants are notified via the block scanning device as shown in Fig. 2B. At the point of notification, participants are required to stop their task, and then press and hold an on-screen button for a fixed



(A)

(B)



(C)

Figure 3. Sequence of human participant assembly task. Search (A), Scan (B) and Assemble (C).

amount of penalty time. Afterward, both team members are allowed to continue their respective tasks. To determine conflict, experiment supervisors operated a panel with three push buttons while monitoring an overhead video feed of the assembly bins. While a participant's hands were over a given bin, the supervisor depressed the corresponding button. The human's location is checked against the robot's location and the robot is halted when the two match.

B. Design

The intent communication modality variable consisted of four conditions based on implementations of motion-based and display-based intent communication. The four conditions were: (1) a control condition, in which neither motion-based or display-based intent communication were used, (2) a motion-based condition, in which only motion-based intent communication was used, (3) a display-based condition, in which only display-based intent communication was used, and (4) a multimodal condition in which both motion-based and display-based intent communication was used in parallel.

Motion-based intent communication is implemented as predetermined motion plans (Fig. 4) based on legible motion that communicated the robot's intent and illegible motion that purposefully concealed the robot's intent. The predetermined legible and illegible motions to each of the three assembly bins are shown in . Each motion has three phases: (1) move-to-box which took 8 seconds, (2) place-in-box which took 7 seconds, and (3) move-from-box which took 7 seconds. Legible motion is planned to move straight to the back of each assembly bin before entering. This legible motion path allows participants to quickly observe the robots intent (goal bin) with no ambiguity. Illegible motion was planned to move to an ambiguous waypoint between two assembly bins before entering the goal bin. This illegible motion path conceals the intent of the robot between to assembly bins for the 8 seconds until the robot enters its goal bin.

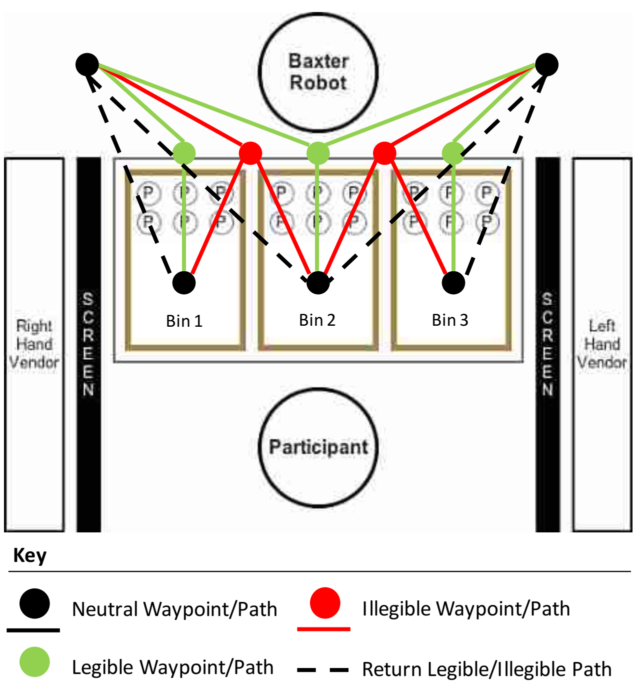


Figure 4. Legible and illegible motion paths within the mock workspace

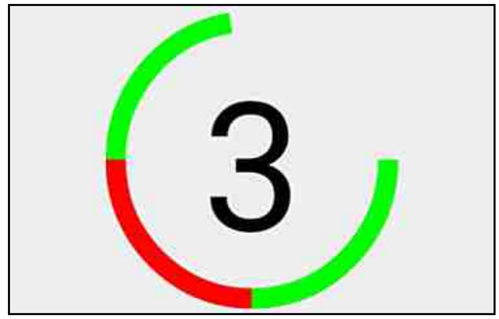


Figure 5. Visual intent communication display

Display-based intent communication was implemented as a visual and auditory display consisting of a visual intent display (Fig. 5) and auditory prompts. The visual intent display consisted of a circular progress bar that counts down the safe (green segments moving to and from the box) and unsafe (red segment, time inside the box when conflict can occur) progress of the robot's action. The visual display also notifies participants of a collision by the circular progress bar turning completely red. The auditory prompt announces which assembly bin the robot is moving to at the start of each action. To an attentive observer, from the start of each action, display-based intent communication will remove ambiguity in the robot's goal assembly bin and the portion of time that that bin will be safe or unsafe for the participant to operate in.

The conflict time penalty variable is the amount of time participants must press and hold the conflict release button on the block scanning device after a conflict occurs. There were two penalty conditions: (1) high penalty which lasted 60 seconds and (2) low penalty that lasted 5 seconds. This variable served to test Hypothesis 2 by contrasting the severity of negative consequences assessed for failing to coordinate. Penalty times were selected based on pilot testing to determine values that would cause either negligible (low) or significant (high) frustration to participants.

C. Dependent Measures

To measure the ability of the implemented intent communication techniques to persistently enable fluent coordination, we used four measures of team fluency across each of the trials based on Hoffman et al. [6].

The first measure recorded was participant idle time. This metric indicates the ability of the intent communication techniques to enable participants to judge when it is safe to operate and hence minimize the amount of available safe operation time wasted by participants in completing their assembly task. Idle time is computed as the total completion time for each trial minus the time that either the robot and participant spent actively working in the assembly bins.

The second measure recorded was the number of conflicts that occurred between the robot and the human in each 12-cycle trial. This measure indicates the ability of intent communication to enable participants to avoid conflict when coordinating with their robotic teammate.

The third and fourth measures recorded were subjective responses to a five-question survey administered post-trial that aimed to gauge subjective fluency of team operations (Q1 and Q2) and subjective transparency of the robotic team member (Q3, Q4 and Q5). The survey questions were

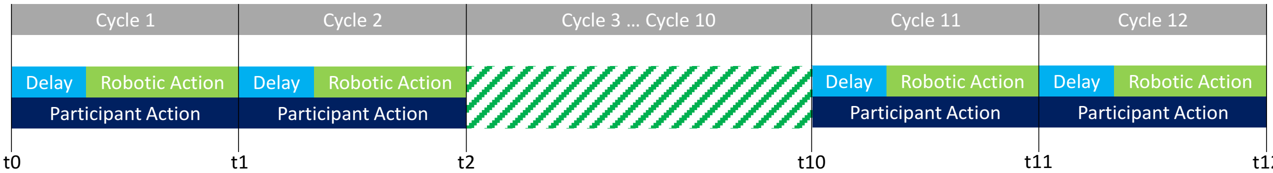


Figure 6. Timeline of cycles and actions taken by each agent. Robot actions are delayed randomly to lessen learning effects

answered on a five-point Likert scale, with Q1 rated on the scale Very Negative to Very Positive, and Q2-5 rated on the scale Strongly Disagree to Strongly Agree. The questions are as follows:

1. What was your overall feeling toward your robotic teammate throughout the trial?
2. The robot and I worked as an effective team.
3. The robot's behavior was confusing throughout the trial.
4. The robot effectively communicated its target boxes throughout the trial.
5. I always understood to which box the robot was moving to throughout the trial.

The first two measures are automatically recorded by the control software and the third and fourth measures were obtained from the post-trial survey administered via the block scanning device.

D. Procedure

First, participants are asked to read and sign an informed consent from and answer a short demographic survey about their age, gender and prior exposure to robots in the workplace.

Next, participants complete an interactive video tutorial that walks them through the rules and a complete overview of the experimental environment including the workspace, intent displays, the participant's task, the robot's task, and the concept and consequences of conflict. Once complete, participants are moved to the workspace and asked to complete two two-block demonstration trials that demonstrated the tasks, the intent displays and a conflict scenario. Then, participants are allowed to ask any last questions they have before beginning the experiment.

Participants complete each of the eight conditions in a unique randomized order to reduce effects of learning or fatigue. To begin each trial, the experiment supervisor instructs participants of whether the penalty was 5 or 60 seconds. The start of each trial is denoted by a 3 second countdown signified by three warning tones and a trial start tone. Participants and the robot each conduct in sequence the twelve assembly and fetching cycles. Once both the participant and robot finish all cycles, a tone sounds to signify the end of the trial. The block scanning device then prompts participants to answer the five post-trial questions. Once completed, the supervisor resets the workspace by removing blocks from assembly pegs and ensuring there are the correct 20 blocks in each bin. Each experiment took around three hours and participants were compensated $15 per hour.

IV. Results

A total of 29 participants were recruited from the local community, 17 of which were male and 12 of which were female between 18 and 61 years of age (M = 26.55, SD = 6.40), all of whom were screened to have normal or corrected to normal vision.

A. Number of Conflicts and Participant Idle Time

A direct comparison of means under each display condition for both participant idle time and number of conflicts (Fig. 7A and Fig. 7B), shows that the multimodal approach and both unimodal approaches yielded a smaller participant idle time and fewer number of conflicts in both high and low conflict penalty conditions when compared to the control condition. By inspection of the plots, idle time was similar between high and low penalty conditions, but a greater number of conflicts was found under low penalty conditions versus high penalty conditions. It also appears that multimodal intent communication performed best, with the lowest mean idle time and lowest mean number of conflicts.

To assess the statistical significance of these differences a two-way repeated measures analysis of variance (ANOVA) was conducted on participant idle time. Inspection of Q-Q plots and the Shapiro-Wilk test suggests that normality was violated. A square root transform of the data improved normality and hence tests are reported for both the original and transformed data. No significant effect of conflict penalties was found for both the original data ($F(1, 28) = 1.189$, $p = 0.281$, $\eta^2_{partial} = 0.041$) or the transformed data ($F(1, 28) = 1.189$, $p = 0.285$, $\eta^2_{partial} = 0.041$). A significant effect of intent communication modality was found for both the original data ($F(3, 84) = 7.111$, $p < 0.001$, $\eta^2_{partial} = 0.203$)

TABLE 1. PAIRWISE BONFERRONI ADJUSTED MAIN EFFECT COMPARISONS OF INTENT COMMUNICATION VS NO-COMMUNICATION CONDITIONS. (SIGNIFICANT AT P < 0.05)

Intent Communication Comparison	**Participant Idle Time**				**Number of Conflicts**		**Subjective Team Fluency**		**Subjective Teammate Transparency**	
	Original Data		Square Root Transform		Original Data		Original Data		Original Data	
	% Change	Sig	% Change	Sig	% Change	Sig	% Change	Sig	% Change	Sig
Control Vs. Multimodal	**Δ = 11.37%**	**p < 0.001**	**Δ = 5.85%**	**p < 0.001**	**Δ = 57.14%**	**p = 0.003**	**Δ = 22.71%**	**p = 0.004**	**Δ = 100.52%**	**p < 0.001**
Control Vs. Display	**Δ = 8.69%**	**p = 0.023**	**Δ = 4.44%**	**p = 0.022**	Δ = 38.46%	p = 0.248	**Δ = 23.59%**	**p = 0.006**	**Δ = 97.13%**	**p < 0.001**
Control Vs. Motion	Δ = 3.80%	p = 0.965	Δ = 1.94%	p = 0.835	**Δ = 47.25%**	**p = 0.001**	Δ = 9.44%	p = 0.529	**Δ = 27.94%**	**p = 0.007**
Multimodal Vs. Display	Δ = -3.02%	p = 1.000	Δ = -1.50%	p = 1.000	Δ = -43.58%	p = 0.586	Δ = -0.72%	p = 1.000	Δ = -1.69%	p = 1.000
Multimodal Vs. Motion	**Δ = -8.54%**	**p = 0.005**	**Δ = -4.16%**	**p = 0.007**	Δ = -23.07%	p = 1.000	**Δ = -10.82%**	**p = 0.041**	**Δ = -36.20%**	**p < 0.001**
Display Vs. Motion	Δ = 5.36%	p = 0.953	Δ = 2.62%	p = 1.000	Δ = -14.28%	p = 1.000	Δ = -11.45%	p = 0.057	**Δ = -35.10%**	**p < 0.001**

and the transformed data ($F(3, 84) = 6.945$, $p < 0.001$, $\eta^2_{partial} = 0.199$). No significant interaction effect was found between conflict penalties and intent communication modality on the both the original ($F(3, 84) = 0.966$, $p = 0.413$, $\eta^2_{partial} = 0.033$) or transformed data ($F(3, 84) = 0.938$, $p = 0.426$, $\eta^2_{partial} = 0.032$). Post-hoc simple effects within each penalty condition show a significant effect of intent communication modality under both high conflict penalties ($F(3, 26) = 9.702$, $p < 0.001$, $\eta^2_{partial} = 0.528$) and low conflict penalties ($F(3, 26) = 3.684$, $p = 0.025$, $\eta^2_{partial} = 0.298$).

To assess significance a two-way repeated measures analysis of variance was conducted on the number of conflicts. Inspection of Q-Q plots and the Shapiro-Wilk test suggests that normality was violated. Multiple data transformations were investigated with minimal improvements in normality, hence we report tests with only original data. A significant effect of conflict penalties was found ($F(1, 28) = 4.802$, $p = 0.037$, $\eta^2_{partial} = 0.146$). In analyzing the effect of intent communication modality Mauchly's test revealed the assumption of sphericity was violated ($\chi^2(5) = 22.011$, $p = 0.001$) and hence the degrees of freedom were corrected using the Greenhouse-Geisser correction with $\varepsilon = 0.632$ (as Greenhouse-Geisser $\varepsilon < 0.75$) yielding a significant effect ($F(1.895, 53.066) = 5.483$, $p = 0.008$, $\eta^2_{partial} = 0.164$). No significant interaction effect was found between conflict penalties and intent communication modality ($F(3, 84) = 0.058$, $p = 0.981$, $\eta^2_{partial} = 0.002$). Post-hoc simple effects within each penalty condition show a significant effect of modality under high penalties ($F(3, 26) = 3.942$, $p < 0.019$, $\eta^2_{partial} = 0.313$) and a marginally significant effect under low penalties ($F(3, 26) = 2.579$, $p = 0.075$, $\eta^2_{partial} = 0.229$).

Pairwise comparisons using the Bonferroni correction (TABLE 1) show a significant decrease in participant idle time and number of conflicts in comparison to the control condition when multimodal intent communication is used. Display-based intent communication significantly improved participant idle time but not the number of conflicts when compared with the control condition. Conversely, motion-based intent communication significantly reduced the number of conflicts, but failed to significantly reduce participant idle time in comparison to the control condition. Multimodal intent communication significantly improved participant idle time but not the number of conflicts when compared to display-based. Multimodal did not significantly improve idle time or number of conflicts in comparison to display-based.

B. Post-Trial Survey

A direct comparison of means for subjective team fluency (Q1 and Q2) and subjective teammate transparency (Q3, Q4, and Q5) within each display condition (Fig. 7C and Fig. 7D),

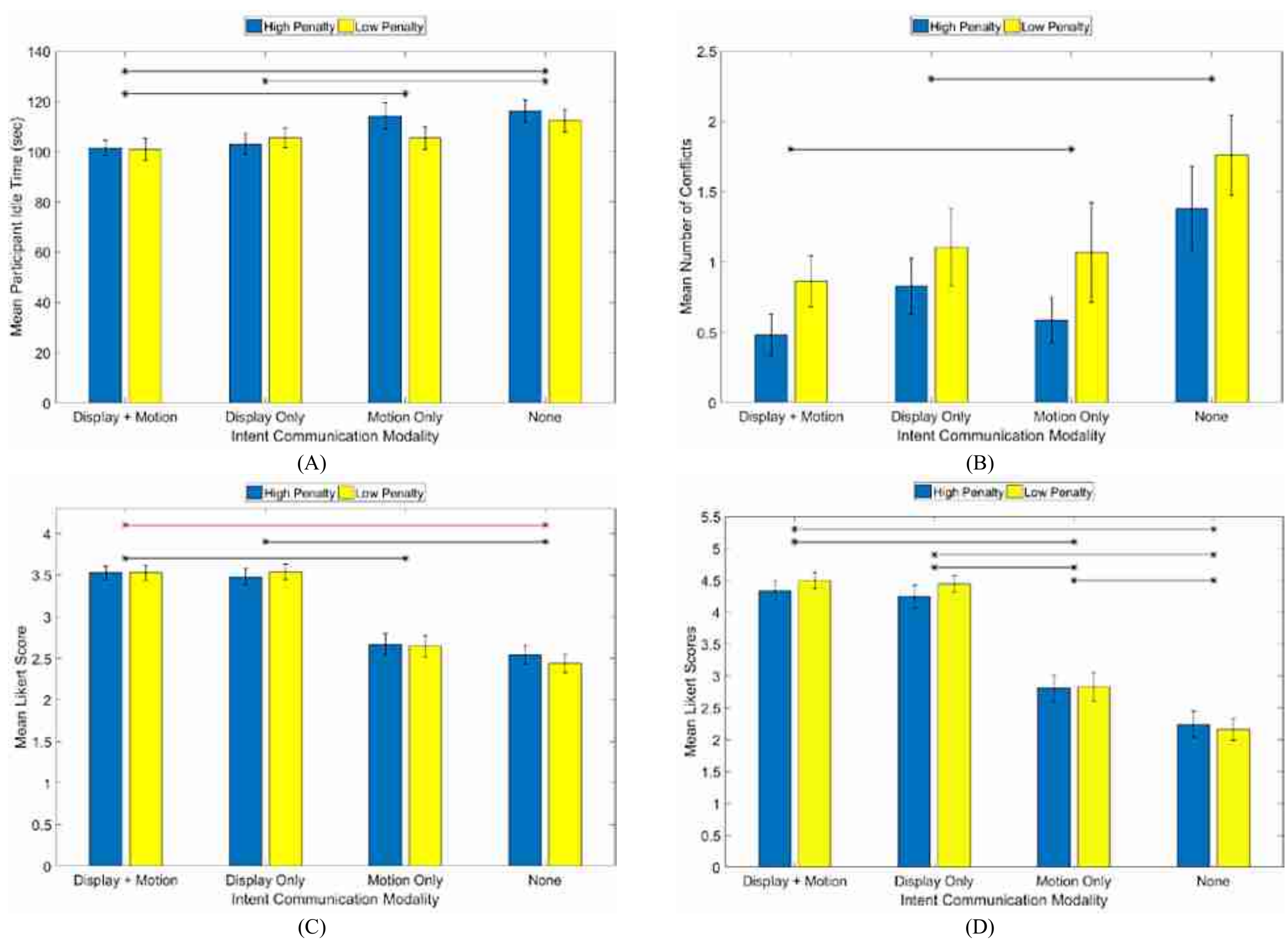


Figure 7. Plots of objective and subjective measures, (A) mean participant idle time, (B) mean number of conflicts, (C) mean Likert scores for subjective team fluency (Q1 and Q2), and (D) mean Likert scores for subjective teammate transparency (Q3, Q4 and Q5). Error bars indicate standard error of the mean. Asterisk lines indicate significant pairwise comparisons between intent communication modalities.

shows that in general, participants report better team fluency and better teammate transparency when multimodal or unimodal intent communication was used. By inspection, subjective ratings of both teammate transparency and team fluency were higher under the low penalty condition in comparison to the high penalty condition. Further, the plots reveal that multimodal intent communication resulted in the highest ratings of teammate transparency and team fluency, although display-based performed similarly well.

A two-way repeated measures analysis of variance was conducted on the mean Likert scores across Q1 and Q2 representing subjective team fluency (Fig. 7C). A significant effect of conflict penalties was found (F(1, 28) = 9.8, p = 0.004, $\eta^2_{partial}$ = 0.259). A significant effect of intent communication modality was found (F(3, 84) = 9.370, p < 0.001, $\eta^2_{partial}$ = 0.251). No significant interaction effect was found between conflict penalties and intent communication modality (F(3, 84) = 0.145, p = 0.932, $\eta^2_{partial}$ = 0.005). Post-hoc multivariate simple effects within each penalty condition show a significant effect of modality under both high penalty (F(3, 26) = 3.278, p = 0.037, $\eta^2_{partial}$ = 0.274) and low penalty conditions (F(3, 26) = 5.167, p = 0.006, $\eta^2_{partial}$ = 0.374).

A two-way repeated measures analysis of variance was conducted on the mean Likert scores across Q3, Q4 and Q5 representing subjective teammate transparency (Fig. 7D). No significant effect of conflict penalties was found (F(1, 28) = 0.924, p = 0.345, $\eta^2_{partial}$ = 0.032). In analyzing the effect of intent communication modality Mauchly's test revealed the assumption of sphericity was violated ($\chi^2(5)$ = 35.046, p < 0.001) and hence the degrees of freedom were corrected using the Greenhouse-Geisser correction with ε = 0.582 (as Greenhouse-Geisser ε < 0.75) yielding a significant effect (F(1.746, 48.879) = 53.986, p < 0.001, $\eta^2_{partial}$ = 0.658). No significant interaction effect was found between conflict penalties and intent communication modality (F(3, 84) = 1.161, p = 0.329, $\eta^2_{partial}$ = 0.040). Post-hoc multivariate simple effects within each penalty condition show a significant effect of modality under both high penalty (F(3, 26) = 19.136, p < 0.001, $\eta^2_{partial}$ = 0.688) and low penalty conditions (F(3, 26) = 34.756, p < 0.001, $\eta^2_{partial}$ = 0.800).

Pairwise comparisons using the Bonferroni correction (TABLE 1) show a significant increase in subjective team fluency and subjective teammate transparency in comparison to the control condition when either multimodal or display-based intent communication is used. Motion-based intent communication had a significant effect on subjective teammate transparency but not subjective team fluency. A multimodal approach significantly improved subjective team fluency and subjective teammate transparency when compared to motion-based intent communication but not display-based intent communication. Display-based intent communication subjectively outperformed a motion-based approach in teammate transparency but not team fluency.

V. Discussion

A. Intent Communication in Pursuit of Coordinated Action

Objective and subjective results show statistically significantly support for Hypothesis 1 as intent communication yielded a greater significant positive effect on participant idle time, number of conflicts, subjective team fluency and subjective teammate transparency. These results provide significant objective and subjective evidence, in line with prior work [8]–[13], [15], that intent communication supports fluent coordination.

Results do not support Hypothesis 2. The significant effect of intent communication on both objective and subjective measures of team fluency did not significantly differ between the high and low penalty conditions. This result surprisingly implies that the positive effect of intent communication was not lessened when the negative consequences for failing to coordinate were smaller. Although this result contradicts our theoretical predictions, it suggests that irrespective of a worker's motivation to engage in coordination-specific activities, intent communication can persistently facilitate fluent coordination. Importantly this confirms the utility of intent communication beyond existing experimental evaluations (e.g. [11], [15]) toward more complex teaming scenarios, in which team members operate concurrently. However, it should be noted that although insignificant, the effect size of intent communication was greater under high penalties than low penalties. Hence further validation of intent communication under scenarios that require concurrent operation should be conducted.

B. Selection of Modalities for Intent Communication

Pairwise comparisons between intent communication modalities showed that in comparison to the control condition a purely display-based approach improved participant idle time, subjective team fluency and subjective teammate transparency, but not the number of conflicts. On the other hand, a purely motion-based approach improved the number of conflicts and subjective teammate transparency but not participant idle time of subjective team fluency when compared to the control condition. Although further experimentation is needed to get a better understanding of this effect we suggest: (1) the pervasive nature of the implemented display-based approach may lead to information overload or inattentional blindness [16], and (2) the act of interpreting intent indirectly through motion may cause a higher cognitive load or feelings of annoyance. Both of these phenomena have the potential to limit a team's ability to fluently coordinate.

The multimodal approach yielded significantly better subjective team fluency and subjective teammate transparency in comparison to purely motion-based communication but not purely display-based communication. Hence, our data only provides mixed support for Hypothesis 3. However, the multimodal approach yielded a significant positive effect on all four measures of team fluency when compared to the control condition. This suggests that although we cannot accept Hypothesis 3, merging both display-based and motion-based intent communication combines the strengths of both unimodal approaches.

A further interesting result is the significant improvement in subjective teammate transparency, and the marginally significant improvement in subjective team fluency when using display-based intent communication. This effect was not seen using motion-based communication. This result suggests that participants feel more confident in the team's ability to coordinate when intent is communicated explicitly.

VI. Conclusion

In this work, we presented an experimental investigation to address two open questions in the use of intent communication toward facilitating fluent coordination in human-robot teams. First, we aimed to evaluate the effectiveness of intent communication in facilitating fluent coordination in scenarios where human team members operate concurrently, alternating between individual and collaborative actions. Second, we aimed to identify the modality which best communicates intent information in scenarios where the resources of human team members are divided between individual and collaborative actions. In addressing these problems, we evaluated two unimodal approaches that use either motion-based or display-based intent communication and a combined multimodal approach in terms of objective and subjective measures of team fluency.

Our investigation presents statistically significant evidence, in line with recent work [8]–[13], [15], that intent communication can facilitate fluent coordination in concurrent team operations. However, our experimental investigation also highlighted shortcomings of a unimodal approach to intent communication, since a purely display-based approach failed to reduce the number of conflicts and a purely motion-based approach failed to significantly improve participant idle time or subjective team fluency. Further, results showed that multi-modal intent communication seems to adopt the strengths of both communication styles, significantly outperforming the control condition in both objective and subjective metrics.

Interestingly, our work also suggests that in general, intent communication is able to persistently facilitate fluent coordination, even in scenarios where humans may not be motivated to engage in coordinated activities. This finding builds upon existing evaluations (e.g. [11], [15]), confirming the efficacy of intent communication in enabling fluent coordination beyond sequential operations toward more complex teaming scenarios that require team members to concurrently work toward the team's goal.

Future work will explore more intent communication modalities (e.g. tactile, speech etc.) as well as the utility of multimodal intent communication in less structured environments. Further, a dynamic model of intent communication might allow robotic team members to dynamically select the most appropriate intent communication modality as the team evolves.

Acknowledgment

This research was supported by grants from the National Science Foundation (No. 1548417 and No. 1503177).

References

[1] J. E. Mathieu, T. S. Heffner, G. F. Goodwin, E. Salas, and J. A. Cannon-Bowers, "The influence of shared mental models on team process and performance.," *J. Appl. Psychol.*, vol. 85, no. 2, pp. 273–83, 2000.

[2] B.-C. Lim and K. J. Klein, "Team mental models and team performance: A field study of the effects of team mental model similarity and accuracy," *J. Organ. Behav.*, vol. 27, no. September 2005, pp. 403–418, 2006.

[3] M. A. Goodrich and A. C. Schultz, "Human-Robot Interaction: A Survey," *Found. Trends® Human-Computer Interact.*, vol. 1, no. 3, pp. 203–275, 2007.

[4] G. Hoffman and C. Breazeal, "Collaboration in Human-Robot Teams," in *AIAA 1st Intelligent Systems Technical Conference*, 2004.

[5] V. Groom and C. Nass, "Can robots be teammates?: Benchmarks in human–robot teams," *Interact. Stud.*, vol. 8, no. 3, pp. 483–500, 2007.

[6] G. Hoffman, "Evaluating Fluency in Human-Robot Collaboration," *HRI Work. Hum. Robot Collab. 2013.*, 2013.

[7] D. R. Olsen and M. A. Goodrich, ""Metrics for Evaluating Human-Robot Interaction," *Proc. Permis*, pp. 1–8, 2003.

[8] M. Williams, S. Abidi, P. Gärdenfors, X. Wang, B. Kuipers, and B. Johnston, "Interpreting Robot Pointing Behavior," 2013, pp. 148–159.

[9] R. M. Holladay, A. D. Dragan, and S. S. Srinivasa, "Legible robot pointing," in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 2014, vol. 2014–Octob, no. October, pp. 217–223.

[10] A. D. Dragan, K. C. T. Lee, and S. S. Srinivasa, "Legibility and predictability of robot motion," *ACM/IEEE Int. Conf. Human-Robot Interact.*, pp. 301–308, 2013.

[11] A. D. Dragan, S. Bauman, J. Forlizzi, and S. S. Srinivasa, "Effects of Robot Motion on Human-Robot Collaboration," *Proc. Tenth Annu. ACM/IEEE Int. Conf. Human-Robot Interact. - HRI '15*, vol. 1, pp. 51–58, 2015.

[12] M. Clamann, M. C. Aubert, and M. L. Cummings, "Evaluation of Vehicle-to-Pedestrian Communication Displays for Autonomous Vehicles," in *2017 96th Annual Transportation Research Board Meeting, Washington, D.C.*, 2017.

[13] O. Ogorodnikova, "Creating an active awareness system for humans in robotic workcell," *Acta Polytech. Hungarica*, vol. 5, no. 2, pp. 113–126, 2008.

[14] M. Kwon, S. H. Huang, and A. D. Dragan, "Expressing Robot Incapability," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction - HRI '18*, 2018, pp. 87–95.

[15] K. Baraka, S. Rosenthal, and M. Veloso, "Enhancing human understanding of a mobile robot's state and actions using expressive lights," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2016, pp. 652–657.

[16] C. D. Wickens, "Multiple Resources and Mental Workload," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 50, no. 3, pp. 449–455, Jun. 2008.

[17] J. MacMillan, E. Entin, and D. Serfaty, "Communication overhead: The hidden cost of team cognition," *Team Cogn. Process*, 2004.

[18] J. A. Cannon-Bowers, E. Salas, and S. A. Converse, "Shared mental models in expert team decision making," in *Individual and group decision making: current issues*, N. . Castellan, Ed. Mahwah: L. Erlbaum Associates, 1993, pp. 221–245.

[19] E. E. Entin and D. Serfaty, "Adaptive Team Coordination," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 41, no. 2, pp. 312–325, Jun. 1999.

[20] J. Langan-Fox, J. Anglim, and J. R. Wilson, "Mental models, team mental models, and performance: Process, development, and future directions," *Hum. Factors Ergon. Manuf.*, vol. 14, no. 4, pp. 331–352, 2004.

[21] L. L. Levesque, J. M. Wilson, and D. R. Wholey, "Cognitive divergence and shared mental models in software development project teams," *J. Organ. Behav.*, vol. 22, no. 2, pp. 135–144, 2001.

[22] E. Salas, N. J. Cooke, and M. a Rosen, "On Teams, Teamwork, and Team Performance: Discoveries and Developments," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 50, no. 3, pp. 540–547, Jun. 2008.

[23] M. Fishbein and I. Ajzen, *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Reading, MA: Addison-Wesley.* Addison-Wesley Pub. Co, 1975.

[24] I. Ajzen, "The theory of planned behavior," *Organ. Behav. Hum. Decis. Process.*, vol. 50, no. 2, pp. 179–211, Dec. 1991.

[25] H. Karvonen and L. Aaltonen, "Intent communication of highly autonomous robots," in *12th ACM / IEEE International Conference on Human-Robot Interaction Workshop on "The Role of Intentions in Human-Robot Interaction,"* 2015.

[26] S. Sheikholeslami, Aj. Moon, and E. A. Croft, "Cooperative gestures for industry: Exploring the efficacy of robot hand configurations in expression of instructional gestures for human–robot interaction," *Int. J. Rob. Res.*, vol. 36, no. 5–7, pp. 699–720, Jun. 2017.

[27] B. Gleeson, K. MacLean, E. Croft, and J. Alcazar, "Gestures for Industry," *Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, pp. 349–356, 2013.