Decision-Making Authority, Team Efficiency and Human Worker Satisfaction in Mixed Human-Robot Teams

Matthew C. Gombolay, Reymundo A. Gutierrez, Giancarlo F. Sturla, and Julie A. Shah
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
gombolay@csail.mit.edu, {ragtz,sturlag}@mit.edu, julie_a_shah@csail.mit.edu

Abstract— In manufacturing, advanced robotic technology has opened up the possibility of integrating highly autonomous mobile robots into human teams. However, with this capability comes the issue of how to maximize both team efficiency and the desire of human team members to work with robotic counterparts. We hypothesized that giving workers partial decision-making authority over a task allocation process for the scheduling of work would achieve such a maximization, and conducted an experiment on human subjects to test this hypothesis. We found that an autonomous robot can outperform a worker in the execution of part or all of the task allocation ($p < 0.001$ for both). However, rather than finding an ideal balance of control authority to maximize worker satisfaction, we observed that workers preferred to give control authority to the robot ($p < 0.001$). Our results indicate that workers prefer to be part of an efficient team rather than have a role in the scheduling process, if maintaining such a role decreases their efficiency. These results provide guidance for the successful introduction of semi-autonomous robots into human teams.

I. INTRODUCTION

In manufacturing, there is a growing desire to leverage the unique strengths of humans and robots to form highly effective human-robot teams [9], [26], [27]. Often robots are not capable of performing the same tasks as their human counterparts, and human workers shift toward the performance of a smaller set of tasks better suited for human dexterity and intelligence, upon the introduction of a robot worker to their environment. The proper functioning of a human-robot manufacturing team requires strict coordination between human and robotic work that satisfies hard temporal and spatial constraints. Academic researchers and industry practitioners alike have developed systems for the planning or scheduling of human-robot work where the humans are either included in the decision-making process [9], [11], [32] or the work is scheduled autonomously [2], [5]. In this work we experimentally investigate whether the robot or the human worker should individually maintain or share decision-making authority over how best to allocate the work in order to maximize both human-robot team fluency and the satisfaction of the human worker.

Human workers often develop a sense of identity and security from their roles or jobs in a factory, and many are used to having some degree of autonomy in decision-making. As a result, a human worker who is tasked by an automated scheduling algorithm may feel devalued. Even if the algorithm increases process efficiency at first, taking control away from human workers may alienate them and, in turn, ultimately damage overall productivity. On the other hand, workers may find the process of scheduling to be burdensome, and prefer to be part of an efficient team rather than have a role in the scheduling process, if maintaining such a role decreases their efficiency. While autonomous scheduling algorithms can provide near-optimal schedules within seconds, we also want to determine how much decision-making authority humans should have in the task allocation process, so that they feel appreciated while still maintaining a high level of team efficiency.

We hypothesize that there is an inverse relationship between human-robot team efficiency and the satisfaction of the human workers, according to how much control the human workers have over *task allocation*, or the assignment of which worker will perform which task. We conducted a human-subject experiment ($n = 24$) where subjects worked alongside both another human and a robot. Rather than finding a desired middle ground between fully autonomous and manual scheduling scenarios, we found statistically significant evidence that giving human subjects more control authority over task allocation negatively influenced team fluency ($p < .02$) and the desire of the subject to work with the robot again ($p < 0.001$). We also found evidence of a complex relationship between human/robot decision-making authority and human preferences over task allocation; people sought looser couplings between human and robot work when they did not retain primary decision-making authority.

II. BACKGROUND

The development of effective human-machine systems has been the focus of research for many in the fields of human factors, robotics, manufacturing and aerospace, to name a few. A key goal of this work has been to leverage the unique strengths of both the human and robot. Researchers have defined a good robot teammate as one that permits teammates to choose actions and timings on the fly, dynamically anticipates and adapts, ensures time-critical tasks are accomplished first and supports interaction that is fluid and natural to the human [6], [12], [23], [29], [31].

The human-robot interface has long been identified as a major bottleneck for the utilization of these robotic systems to their full potential [7]. As a result, significant research

efforts have been aimed at easing the use of these systems in the field, including the careful design and validation of supervisory and control interfaces [4], [10], [15], [21], [20].

Related research efforts have focused on the inclusion of a human in the decision-making loop to improve the quality of task plans and schedules for robots or semi-autonomous systems [9], [10], [11]. This is particularly important if the human operators have knowledge of factors not explicitly captured by the system model or if scheduling decisions have life and death consequences. In a study of aircraft carrier flight deck operations, veteran operators used heuristics to quickly generate an efficient plan and outperformed optimization algorithms [28]. Other works aimed to leverage the strengths of both humans and machines in scheduling by soliciting user input in the form of quantitative, qualitative, hard or soft constraints over various scheduling options. Recommended schedules were then autonomously compiled and provided to users [3], [16], [17], [22], [32]. While developed to support scheduling of human teams, these approaches can be readily extended to human-robot teams.

Supervisory systems have also been developed to assist human operators in the coordination of the activities of either four-robot or eight-robot teams [8]. Experiments demonstrated that operators were less able to detect key surveillance targets when controlling a larger number of robots. Similarly, other studies have investigated the perceived workload and performance of subjects operating multiple ground mobile-based robots [1]. Findings indicated that a number of robots greater than two greatly increased the perceived workload and decreased the performance of the human subjects.

Some industry practitioners, however, have taken a different approach to the scheduling human-robot teams. When fulfilling online orders in warehouses, workers have to navigate the warehouse to find the correct items, pick up the item and then return to the packaging area to complete the order. Kiva Systems has developed robots that are able to fetch these items for the worker and ensure that each worker is never idle while waiting for the next item to package. One might initially think that narrowing the role of workers in a factory setting may cause them to feel less important; however, CNN has reported that "robots make for a more pleasant work environment" because they "eliminate much of the mundane physical labor employees once did to retrieve products off shelves." [2].

In this work, we are motivated by the application of robotics in the manufacturing domain, where human workers will perform physical tasks in coordination with robotic partners. In some cases, the human workers may also be responsible for tasking the team and tracking progress. We seek to understand how much control human workers should have over the assignment of roles and schedules when working in teams with robots. The following sections will describe our experiment to lend insight into the relationship between team efficiency and worker satisfaction, as a function of the control authority of human workers over team scheduling.

III. AIM OF THE EXPERIMENT

We sought to understand the contributions of efficiency, worker decision-making authority and human idle time to objective and subjective measures of team performance and worker satisfaction. Understanding the relationship between these measures will provide researchers and industry practitioners with better insight into how to design successful human-robot teams.

A. Independent Variables

In our experiment, we control the level of decision-making authority over task allocation the worker has during work scheduling for their team. This independent variable can have one of three values:
1) *Manual Control*-The subject decides who will perform which tasks
2) *Semi-Autonomous Control*-The subject decides which tasks he will perform, while the robot allocates the remaining tasks to itself and the human assistant
3) *Autonomous Control*-The robot allocates all tasks.

The robot performs task sequencing in all three conditions. We explore the decision-making authority over task allocation alone, rather than over both task allocation and sequencing, in order to isolate the effects of task allocation and mitigate experimental confound. We leave investigation of sequencing and joint task allocation and sequencing to our future experimentation.

B. Hypotheses

H1 *Team productivity degrades when the subject has more control over the rescheduling process. As a metric of productivity, we measure both the time it takes to reschedule and the time it takes to finish all tasks.*
Determining the optimal schedule while under hard upper- and lower-bound temporal constraints is NP-Hard [5]. Even for problems of a modest size, optimal scheduling becomes intractable in these circumstances. While we have seen a great deal of work in the development of supervisory control interfaces and human in-the-loop systems to leverage the strengths of human insight and the computational power of autonomous scheduling algorithms ([4], [15], [21], [9], [11]), we expect a near-optimal scheduling algorithm to generate better schedules than those generated by the human subjects.
H2 *Subjects prefer having partial control over the rescheduling process rather than complete control, and prefer having complete control to having no control. We utilize a series of subjective Likert-scale questions to determine which level of control the subjects prefer.*
We posit that allowing subjects to decide which tasks they will perform and having the robot complete the remainder of the rescheduling will be most satisfying for the subject. In this scenario, a subject can select their preferred tasks according to perceived physical and mental demands, and has a more substantial role in the success of their team. We believe that both of these factors will contribute to the subject being most satisfied by sharing control of the rescheduling process with the robot. On the other hand, giving subjects

the responsibility of quickly and optimally rescheduling all work will be overwhelming and least desirable. Furthermore, allowing the robot to have complete control will improve team fluency, but at the cost of possibly devaluing the role of the subject.

H3 *Subjects are more satisfied with their experience working on the team when they are less idle. To test this hypothesis, we utilize timing information of task execution during the assembly process and the same set of subjective Likert-scale questions used to test Hypothesis 2.*

Many studies have used idle time as a proxy for team fluency [19], [25], [30], and we posit that subjects' satisfaction is negatively correlated with idle time.

IV. EXPERIMENTAL METHODS

We designed an environment analogous to a manufacturing setting. The subject is a member of a human-robot team responsible for completing a set of tasks that includes both the fetching and assembly of part kits. For each trial, the team must schedule and complete this set of tasks. In our experiment, the goal was to assemble various components of a Lego kit, as shown in Figure 1. A video describing the experiment can be found at `http://tiny.cc/k4hzgx`.

A. Materials and Setup

We used a Willow Garage PR2 platform as shown in Figure 2, as the robotic assistant for our human-robot team. Relevant to the experiment, the PR2 has a holonomic base with optical encoders for each wheel and a $270\deg$ Hokuyo laser at the base. We mapped the laboratory using this laser, as well as the standard *Gmapping* package in the Robot Operating System (ROS). For navigation, we used the Adaptive Monte Carlo Localization (AMCL) [13] probabilistic localization package and a hybrid-dynamical *proportional-derivative* (PD) controller. The locations of the pick-up and drop-off locations for each part kit were hard-coded into the robot controller. The inspection component of the fetching task was simulated.

B. Human-Robot Team Composition

Our human-robot manufacturing team consisted of the human subject, a robotic assistant and a human assistant. The human subject was capable of both fetching and building, and the robot assistant was only capable of fetching. One of the experimenters played the role of a third teammate for all subjects and was capable of both fetching and
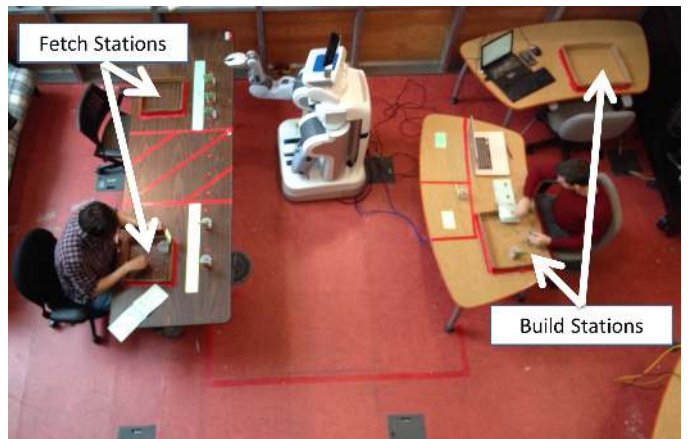

Fig. 2. This figure depicts a diagram of the laboratory room where the experiment took place. There are two locations where the human and robot workers can inspect part kits during a fetching task, and two locations where the human workers can build the part kits.

building. The third human teammate was included to more realistically represent the composition of a human-robot team in a manufacturing setting. The human subjects either performed the task allocation or shared the decision-making authority over task allocation with the robotic teammate, depending on the experimental condition. The robot assistant was always responsible for sequencing the team's work. The third teammate did not provide any decision-making assistance to the subject or the robot.

C. Experiment Task

In our scenario, the fetching of a kit required walking to one of two inspection stations where the kits were located, inspecting the part kit and carrying it to the build area. The architecture of our fetching task is analogous to what is required in many manufacturing domains: To adhere to strict quality assurance standards, fetching a part kit requires verification from one to two people that all correct parts are in the kit, and certification from another person that the kit has been verified.

There were a number of constraints imposed on the analog assembly process, in order to model relevant constraints encountered during assembly manufacturing: First, a part kit must have been fetched before it could be built. Also, no two agents were able to occupy the same fetching or build station. There were two fetching and two build stations, as shown in Figure 2. Four part kits were located at one fetching station, and four kits were located at the second fetching station. When fetching a part kit, inspection of that kit must have occurred at the fetching station where it was initially located.

Because there were an equal number of building stations and agents able to build, there were no additional constraints imposed exclusively on build tasks. However, because there were three agents who could fetch and only two fetching stations, the agents were required to take turns using the

Fig. 1. The Assembled Lego Model.

fetching stations. Allowing workers to sort through parts from multiple kits at the same location risked mixing the wrong part with the wrong kit. We imposed a 10-minute deadline from the time that the fetching of a part kit began until that part kit had been built, for similar reasons. In manufacturing, if a part or part kit is missing from an expected location for too long, work in that area of the factory will temporarily halt until the missing pieces are found.

D. Formulation of the Human-Robot Scheduling Problem

Assembly of the Lego model involved eight tasks $\tau = \{\tau_1, \tau_2, \ldots, \tau_8\}$, each of which was composed of a *fetch* and *build* subtask $\tau_i = \{\tau_i^{fetch}, \tau_i^{build}\}$. The time each subject took to complete each subtask $C_i^{subject-fetch}$ and $C_i^{subject-build}$ was measured during an experiment training round. The timings for the robot $C_i^{robot-fetch}$ and human assistant $C_i^{assist-fetch}$ and $C_i^{assist-build}$ (performed by an experimenter) were collected prior to the experiments. Constraints on lowerbound completion time of tasks are presented in Equations 1-5.

$$f_i^{build} - s_i^{build} \geq C_i^{subject-build} - M(1 - A_{\tau_i^{build}}^{subject}) \quad (1)$$

$$f_i^{fetch} - s_i^{fetch} \geq C_i^{subject-fetch} - M(1 - A_{\tau_i^{fetch}}^{subject}) \quad (2)$$

$$f_i^{fetch} - s_i^{fetch} \geq C_i^{robot-fetch} - M(1 - A_{\tau_i^{fetch}}^{robot}) \quad (3)$$

$$f_i^{build} - s_i^{build} \geq C_i^{assist-build} - M(1 - A_{\tau_i^{build}}^{assist}) \quad (4)$$

$$f_i^{fetch} - s_i^{fetch} \geq C_i^{assist-fetch} - M(1 - A_{\tau_i^{fetch}}^{assist}), \quad (5)$$

where $A_{\tau_i^{subtask}}^{agent}$ is a binary decision variable for the assignment of $agent \in \{subject, robot, assist\}$ to each $subtask \in \{fetch, build\}$ of $\tau_i \in \boldsymbol{\tau}$.

Constraints 6 and 7 also ensured that each agent performed one subtask at a time.

$$s_x^y - f_i^j \geq -M\left(1 - x_{\langle \tau_i^j, \tau_x^y \rangle}\right)$$
$$- M\left(2 - A_{\tau_i^j}^{agent} - A_{\tau_x^y}^{agent}\right), \forall \tau_i^j, \tau_x^y \in \boldsymbol{\tau} \quad (6)$$

$$s_i^j - f_x^y \geq -Mx_{\langle \tau_i^j, \tau_x^y \rangle}$$
$$- M\left(2 - A_{\tau_i^j}^{agent} - A_{\tau_x^y}^{agent}\right), \forall \tau_i^j, \tau_x^y \in \boldsymbol{\tau}, \quad (7)$$

where $x_{\langle \tau_i^j, \tau_x^y \rangle} \in \{0, 1\}$ is a binary decision variable specifying whether τ_i^j comes after or before τ_x^y.

Temporal constraints 8 and 9 ensured that parts for each task were fetched before building began, and that building finished within $D = 10$ minutes of fetching the parts.

$$\infty \geq s_i^{build} - f_i^{fetch} \geq 0, \forall \tau_i \in \boldsymbol{\tau} \quad (8)$$

$$D \geq f_i^{build} - s_i^{fetch} \geq 0, \forall \tau_i \in \boldsymbol{\tau}, \quad (9)$$

where s_i^{build}, s_i^{fetch} and f_i^{build}, f_i^{fetch} are the start and finish times of the build and fetch subtasks, respectively.

The spatial constraint in Equation 10 ensured that no two agents occupied the same fetching station at the same time.

$$s_j^{fetch} - f_i^{fetch} \geq 0 \ \vee \ s_i^{fetch} - f_j^{fetch} \geq 0,$$
$$\forall \tau_i, \tau_j \in \boldsymbol{\tau} \text{ s.t. } R_i^{fetch} = R_j^{fetch}, \quad (10)$$

where R_k^{fetch} denotes the physical floor area reserved for the fetching subtask τ_k^{fetch}. Fetching subtasks $\{\tau_i^{fetch} | i \in \{1, 2, 3, 4\}\}$ used the first inspection station R_1^{fetch}, and $\{\tau_{i'}^{fetch} | i' \in \{5, 6, 7, 8\}\}$ used the second inspection station R_2^{fetch}.

Finally, one pair of tasks $\boldsymbol{\tau_{two-step}} = \langle \tau_3, \tau_4 \rangle$ was related through a precedence constraint. Specifically particants were instructed that the first task, τ_3 in the pair be completed before starting to fetch parts for the second task τ_4. This constraint is presented in Equation 11.

$$s_j^{fetch} - f_i^{build} \geq 0, \forall \langle \tau_i, \tau_j \rangle \in \boldsymbol{\tau_{two-step}} \quad (11)$$

To objective of the problem (Equation 12) was to minimize the maximum amount of work assigned to any one agent, while satisfying the constraints in Equations 1 - 11.

$$obj = \arg\min \left(\max_{agent} \left(\sum_{\tau_i} \left(\sum_{subtask} C_{\tau_i^{subtask}}^{agent} A_{\tau_i^{subtask}}^{agent} \right) \right) \right) \quad (12)$$

E. Scheduling Mechanism

To enable the robot to schedule with varying degrees of decision-making input from the subject, we adapted Tercio, a fast, near-optimal scheduling algorithm that divides the scheduling process into task allocation and sequencing subroutines [14]. The algorithm works by solving for the optimal task allocation and then finding a corresponding task sequence. If the schedule does not satisfy a specified makespan, a second iteration is performed by finding the second-most optimal task allocation and corresponding sequence. The process terminates when the user is satisfied with the schedule quality, or when no better schedule can be found. In this experiment we specified that Tercio run for 25 iterations and return the best schedule.

In the scenario where the subject performed task allocation, the robot used Tercio to sequence tasks and return a flexible, dispatchable schedule [24]. When the subject decided which tasks he or she would perform, the robot used Tercio to find an efficient schedule by iterating over different allocations to the robot and the human assistant. Tercio receives upperbound, lowerbound, and expected duration of each task, and uses the expected durations to compute near-optimal schedules. The upperbound and lowerbound times are used in computing the flexible, dispatchable schedule to allow subjects the flexibility to work faster or slower than expected. We set the lowerbound duration of subtasks assigned to the subject to be 25% faster than their timings during training to mitigate subject idle time due to learning effects.

F. Human-Robot Coordination

Subjects were provided each agent's expected time to complete each of the sixteen subtasks, for experiment conditions where the subject performed the task allocation. In the manual condition, subjects specified the assignment of $agent \in \{subject, robot, assist\}$ to each $subtask \in \{fetch, build\}$ of $\tau_i \in \boldsymbol{\tau}$, by writing the assignment list on a blank paper. The experimenter input the data to the scheduling algorithm, which then optimized the sequencing of subtasks $\{x_{<\tau_i^j, \tau_n^m>}\}$.

In the semi-autonomous condition, the subjects chose only the subtasks that they would complete themselves, and the robot allocated the remaining subtasks. Again the robot sequenced all subtasks $\{x_{<\tau_i^j, \tau_n^m>}\}$. In the autonomous condition, the scheduling algorithm optimized both the allocation and sequencing of subtasks.

The robot served as a central coordinating agent, communicating to the subject and the human teammate when to start their next subtasks. The subject and human teammate confirmed to the robot as they started or finished subtasks. The team members communicated this information by sending simple, text-based messages over a TCP/IP GUI[1].

G. Procedure

We first introduced the subject to the manufacturing scenario. Subjects were told that they were a member of a human-robot manufacturing team. The manufacturer had recently acquired a new robot to work alongside people in the factory to improve the manufacturer's productivity. We explained the various temporal and spatial constraints of our analog manufacturing task as well as the capabilities and roles of each team member.

We then conducted a training round where the subject fetched and built each of the eight part kits. We timed how long it took the participant to complete each task, and provided this information to the robot and the subject for use in scheduling the work. Participants were instructed to work as quickly as possible without making mistakes. Next, the experimenter explained the constraints imposed on the assembly process, and the subject was trained on how to communicate with the robot via the TCP/IP GUI.

We then performed three trials in which the subject was exposed to each of the three conditions (manual, semi-autonomous or autonomous control), varying the order of the conditions across subjects. Each trial consisted of rescheduling the work and completing all tasks according to that schedule. In the scenarios where the subject participated in the task allocation process, we provided the subject with information about how long it took for each agent to perform each task. The subject was instructed to quickly construct an efficient task allocation with the goal of minimizing the sum of the time spent rescheduling and completing the tasks. Subjects took on approximately 5 minutes when asked to

[1]SocketTest v3.0.0 ©2003-2008 Akshathnkumar Shetty (http://sockettest.sourceforge.net/)

allocate all of the tasks, and approximately 2 minutes when deciding which tasks to complete themselves.

Either autonomously or according to the task allocation information provided by the subject, the robot completed the rescheduling and the assembly process began. Each trial took approximately 15 minutes. After each trial, subjects were asked to answer a post-test questionnaire with 21 Likert-scale questions assessing their experience. The experiment concluded with a final post-test questionnaire with three Likert-scale questions and two free-response questions, as shown in Tables I and II.

H. Experimental Design

The goal of the experiment was to understand the relationship between efficiency and worker satisfaction, as a function of how much control the worker has over his or her own role on the team. Our experiment used a within-subjects design, allowing for more powerful statistical testing. In prior experience, we have seen that human subjects build Lego models at vastly different rates. By utilizing a within-subjects design, we mitigated the effects of inter-subject variability. However, because of this design, we needed to account for possible learning effects over the different trials. The speed with which subjects build generally increases with practice. To block for this factor, we balanced the assignment subjects into groups for each of the $k!$ orderings of our $k = 3$ conditions.

To block for variability in the characteristics of the human assistant on the team, a laboratory researcher played the role of the human assistant for all trials. This assistant performed tasks at a nearly constant speed and did not aid the subject in rescheduling the work.

I. Objective Evaluation

Objective measures of team fluency consist of *assembly time*, *rescheduling time* and *idle time*. "Assembly time" is defined as the difference between the time the last task was completed and the time the first task was initiated. "Rescheduling" time is defined as the sum of the time it took the subject to allocate tasks when the subject was involved and the time it took the robot to complete the remainder of the scheduling work. (The experimenter was required to input the task allocation of the subject into the robot's scheduling algorithm, but we did not include this time as a part of the rescheduling time.) Lastly, we defined "idle time" as the sum of the periods during which the subject was not working.

J. Subjective Evaluation

Subjects received post-trial questionnaires after each trial, consisting of 21 Likert-scale questions, as shown in Table I. Hoffman proposed a set of composite measures for the evaluation of human-robot fluency [18]. Questions 1-3 corresponded to Hoffman's measure of *Robot Teammate Traits*, and Questions 4-13 represented Hoffman's adaptation of the "Working Alliance Index" for human-robot teams, measuring the quality of the alliance amongst the teammates. We added questions 14-21 based on our own insight. Subjects were not

TABLE I
Subjective Measures - Post-Trial Questionnaire

Robot Teammate Traits
1. The robot was intelligent.
2. The robot was trustworthy.
3. The robot was committed to the task.
Working Alliance for Human-Robot Teams
4. I feel uncomfortable with the robot. (reverse scale)
5. The robot and I understand each other.
6. I believe the robot likes me.
7. The robot and I respect each other.
8. I feel that the robot worker appreciates me.
9. The robot worker and I trust each other.
10. The robot worker perceives accurately what my goals are.
11. The robot worker does not understand what I am trying to accomplish.
12. The robot worker and I are working towards mutually agreed upon goals.
13. I find what I am doing with the robot worker confusing. (reverse scale)
Additional Measures of Team Fluency
14. I was satisfied by the teams performance.
15. I would work with the robot the next time the tasks were to be completed.
16. The robot increased the productivity of the team.
17. The team collaborated well together.
18. The team performed the tasks in the least time possible. 19. The robot worker was necessary to the successful completion of the tasks.
20. The human worker was necessary to the successful completion of the tasks.
21. I was necessary to the successful completion of the tasks.

TABLE II
Subjective Measures - Post-Test Questionnaire

Overall Preference
22. If it was the PR2s job to reschedule the work, I would want to work with the robot again.
23. If it was my job to reschedule my work and the PR2 reschedule the work for the PR2 and my human teammate, I would want to work with the robot again.
24. If it was my job to reschedule the work for myself, my human teammate, and the PR2, I would want to work with the robot again.
Open Response Questions
25. Which of the three scenarios did you prefer and why?
26. If you were going to add a robotic assistant to a manufacturing team, to whom would you give the job of rescheduling the work and why?

informed of their rescheduling and build times during the experiment.

Subjects also received a post-test questionnaire after completing the three trials. This questionnaire gathered demographic information, and included three additional Likert-scale questions summarizing the experience of the subject, as well as two open-ended questions.

V. RESULTS

In this section, we report the demographics of the participants, as well as statistically significant and insightful findings from our experiment. We define statistical significance at the $\alpha = .05$ level.

A. Participants

Twenty-four participants were included in the experiment. Each participant worked on the human-robot manufacturing team under each level of decision-making authority for a within-subjects design. To control for learning effects, participants were balanced between one of six groups, including one group for each of the six possible sequences of the three conditions and four subjects for each sequence.The participants (14 men and 10 women) had an average age of 27 ± 7 years (minimum and maximum ages were 20 and 42) and were recruited via email and fliers distributed around a university campus.

B. Objective Measures of Human-Robot Team Fluency

We consider the team's assembly time and the subjects' rescheduling time as a function of the subject's decision-making authority. Recall that hypothesis **H1** predicts that the team will be more fluent, in terms of both assembly and rescheduling time, when the robot has more control authority over task allocation. Rescheduling and assembly times are depicted in Figure 3.

Analysis of variance demonstrated statistically significant differences in the distribution of rescheduling time as a function of decision-making authority, $F(2, 69) = 55.1$, $p < 0.01$. Rescheduling time in the autonomous condition ($M = 30$, $SD = 0$) was lower than in the semi-autonomous condition ($M = 108$, $SD = 69$), $t(23) = 7.24$, $p < 0.01$. Likewise, rescheduling time in the semi-autonomous condition was lower than in the manual condition ($M = 315$, $SD = 154$), $t(23) = 7.23$, $p < 0.01$.

Repeated-measure analysis of variances demonstrated significant differences in assembly time, as a function of condition $F(2, 46) = 3.84$, $p = .03$. Assembly time in the autonomous condition ($M = 520$, $SD = 60.6$) was faster than in the semi-autonomous ($M = 564$, $SD = 83.9$), $t(23) = 2.37$, $p = 0.01$, and manual conditions ($M = 582$, $SD = 115$), $t(23) = 2.18$, $p = 0.02$.

Learning effects were found for assembly time as a function of trial number (ANOVA $F(2, 69) = 3.68$, $p = .03$). Specifically, third-trial assembly times ($M = 519$, $SD = 85$) were lower than those in the first ($M = 585$, $SD = 49$), $t(23)$, $p = .002$ and second trials ($M = 567$, $SD = 75$), $t(23)$, $p = .022$. Nonetheless, the k–factorial design counterbalanced the learning effects, and results indicated significant differences in both assembly time, $F(2, 46) = 3.84$, $p = .03$, and rescheduling time, $F(2, 69) = 55.1$, $p < 0.01$ as a function of level of automation.

C. Subjective Measure of Satisfaction as a Function of Task-Allocation Authority

Recall that our second hypothesis **H2** states that workers prefer partial authority over the task process rather than total control, and that having no control is preferable to having complete control.

An omnibus Friedman test confirmed a statistically significant difference in the distribution of a subset of the Likert-scale responses for the three conditions, as shown in Table

Question	Omnibus	Auto. v. Man.	Semi v. Man.	Auto. v. Semi.
1	**$p = .031$**	**$p = .008$**	$p = .073$	$p = .059$
5	**$p = .046$**	**$p = .007$**	$p = .111$	$p = .105$
10	**$p = .009$**	**$p = .001$**	**$p = .008$**	$p = .063$
11	**$p = .002$**	**$p < .001$**	**$p = .002$**	**$p = .017$**
14	**$p = .005$**	**$p < .001$**	**$p = .003$**	**$p = .008$**
15	**$p = .033$**	**$p = .011$**	$p = .064$	$p = .064$
16	**$p = .018$**	**$p = .004$**	**$p = .048$**	**$p = .031$**
18	**$p = .018$**	**$p < .001$**	**$p = .015$**	**$p = .002$**

III. A pair-wise Friedman test confirmed our hypothesis that subjects were more satisfied under the autonomous and semi-autonomous conditions than the manual condition for the questions listed in Table III.

However, there did not exist a single question for which subjects favored the semi-autonomous condition over the autonomous condition. A post-hoc Friedman test with a requisite Bonferroni correction of $\frac{\alpha}{3}$ indicated that subjects were significantly more satisfied with team performance (Question 14, $p = .008$) under the autonomous condition than the semi-autonomous condition. Likewise, subjects agreed more strongly under the autonomous condition that the team performed the tasks within the least amount of time (Question 18, $p = .002$).

The post-test questionnaire included three questions designed to determine whether subjects would be more likely to work with the robot again given the level of decision-making authority allotted to the subject and the robot. Applying the omnibus Friedman test across Questions 22-24 from Table II, we observed a statistically significant difference in subjects' responses to these questions ($p < 0.001$). Post-hoc analysis, using pair-wise comparison with a Bonferroni correction, confirmed that subjects agreed that they were more likely to work with the robot again if the robot performed task allocation autonomously than if the subject and the robot shared task allocation authority ($p < 0.001$) or if the subject had complete task allocation authority ($p < 0.01$). Similarly,
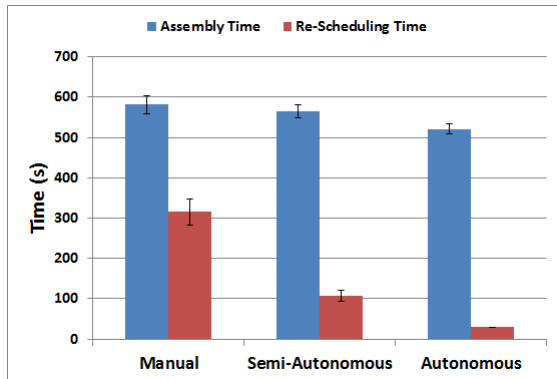
Manual
"People may resent being told what to do by a robot...worker dissatisfaction would probably impact efficiency to the point where any bit of time gained by the robot would be lost by the *demotivation* [sic] of the workers." "There is something soul-sucking about taking the thinking away from the workers"
Semi-autonomous
"I prefer the scenario where I pick the tasks I want because some task are more fun for me than others...even if it might slightly increase completion time." "I got to schedule my work and the robot filled in the rest of the schedule with the purpose of optimizing time"
Autonomous
"It removes the possibility of scheduling being influence by the ego of the team leader." "I never felt like I was wasting time or waiting when the PR2 schedule tasks." "Easier for the robot to deal with scheduling many complex tasks than it is for a human because it can consider all at once [without] getting overwhelmed."

subjects were more likely to report they would work with the robot again if the robot and the human shared task allocation authority than if the subject had sole authority for task allocation ($p < 0.01$).

D. Analysis of Open-Ended Responses

Questions 25 and 26 of the post-trial questionnaire offered subjects the opportunity to provide open-ended responses to prompts on which condition they preferred and to whom they would give control of the task allocation in a manufacturing setting. While the majority of subjects' responses were supportive of a robotic assistant that autonomously allocates work, we also provide representative responses from subjects who preferred the manual, semi-autonomous, and autonomous conditions as shown in Table IV. While most of the subjects' responses directly supported one of the three experimental conditions, some subjects suggested a blended level of control where the robot would "assign tasks but allow person to override (if for example they become overwhelmed or bored)."

E. Subject Idle Time and Satisfaction

Our third hypothesis **H3** states that subjects are more satisfied working on a human-robot team when the they are less idle. Our post-test questionnaire prompted subjects to rate the degree to which they would want to work with the robot again depending on the robots role in the scheduling process (one question for each of the three conditions). Both idle time and subject satisfaction are dependent variables in our experiment. Experiment conditions (autonomous, semi-autonomous, and manual) were ranked according to each subject's preference. The condition's rank was then plotted against the corresponding idle time for each subject-condition pair. The Pearson product-moment correlation coefficient of satisfaction and idle time ($r = 0.125$) was not significant ($t(23) = 0.90$, $p = 0.367$).

Fig. 3. This figure shows the average and standard error for the assembly times in each condition.

F. Posthoc Analysis

We conducted a *posthoc* analysis to better understand differences in the ways people allocate work under the various experiment conditions. In Definition 1, we define a metric, Frequency of Independent Work Allocation (FIWA), to assess the degree to which the agent in charge of task allocation decouples his/her work from the rest of the agents in the team.

Definition 1: Frequency of Independent Work Allocation (FIWA) - the number of instances in which the agent is allocated to complete either:

1) A subtask that does not depend on another subtask being completed before its execution (e.g., τ_1^{fetch})
2) A pair subtasks that are linked by precedence (e.g., $\left\langle \tau_1^{fetch}, \tau_1^{build} \right\rangle$ or $\left\langle \tau_3^{build}, \tau_4^{fetch} \right\rangle$).

We investigated differences in a subject's FIWA score as a function of their decision-making authority. The FIWA score for a subject was greater in the semi-autonomous condition than the autonomous condition 58.3% of the time, but was greater in the manual versus autonomous condition only 37.5% of the time. A z–Test for two proportions demonstrated statistical significance of this difference $z = 1.72$, $p = 0.043$. As such, we establish a new hypothesis for testing in a future experiment: **H4** *In a scenario where the goal is to maximize team efficiency, a human worker is more likely to allow for interdependent work between himself and his team if the subject is responsible in allocating work to the entire team.*

VI. Discussion

A. Guidance on Deploying Autonomous Robot Teammates

The aim of this study was to determine how much control a human member of a human-robot team should have over his or her robot counterpart in order to maximize team efficiency and worker satisfaction. We hypothesized that giving workers some control over the task allocation process would increase satisfaction without too great a sacrifice to team efficiency; however, autonomous control yielded improvements in objective and subjective measures, as compared to manual or semi-autonomous control. This finding is in keeping with anecdotal evidence that subjects prefer working with highly autonomous robots [2].

These results provide guidance for the successful introduction of robots into human teams. First, providing human teammates subjects more decision-making authority over robot behavior is not sufficient to improve the worker satisfaction, and may degrade team performance. Also, team fluency does appear to positively correlate with willingness to collaborate with robotic technology. Second, these experiments provide preliminary evidence that there exists a complex relationship between human/robot decision-making authority and human preferences over task allocation; people may voluntarily seek looser couplings between human and robot work when they do not retain primary decision-making authority. This preference may negatively affect team performance, and warrants investigation in a future study.

B. Limitations and Future Work

There are limitations to our experimental findings. Our sample population consisted of college students and young professionals whose livelihoods are not threatened by the possibility of robots replacing them. Providing manufacturing workers with more control in the decision-making process may still influence the satisfaction of those workers. However, our findings suggest that team fluency is also likely to be an important component in the successful introduction of robot teammates. To better understand the relative contributions of team fluency and decision-making authority towards worker satisfaction in manufacturing, we will conduct a future study where we specifically recruit manufacturing workers.

Each participant in our experiment worked worked with the human-robot for one, 90–minute period. However, manufacturing workers would be working with robots every workday, possibly for years. Human workers may have strong preferences for some jobs over others, and may make different choices or have different preferences in task allocation when working with robots every day for the long term. We propose a longitudinal study is necessary to observe the trajectory of human worker satisfaction over time, since the short and long-term effects of decision-making authority may differ.

VII. Conclusion

With the increasing desire and ability to integrate autonomous robotic agents into manufacturing environments, it is important to understand how much decision-making authority human workers should have over their robotic counterparts when allocating tasks to human and robot team members. While autonomy can improve team efficiency, providing a worker too little or too much control may be alienating or overwhelming, respectively. We conducted an experiment with human subjects to determine how much control a worker should have over the task allocation process. We found that an autonomous robot can outperform a human worker when conducting part ($p < 0.001$) or all of the task allocation ($p < 0.001$). However, rather than finding an ideal balance of control authority to maximize worker satisfaction, we observed that workers preferred to give control authority to the robot ($p < 0.001$). Our results suggest that providing workers with a role in the allocation of tasks to their robotic counterparts may not be an effective method of improving worker satisfaction. Rather, team fluency may more strongly influence worker satisfaction than level of decision-making authority.

REFERENCES

[1] Julie A. Adams. Multiple robot-single human interaction: effects on perceived workload and performance. *Behavior and Information Technology*, 28(2):183–298, 2009.

[2] Jennifer Alsever. Robot workers take over warehouses. *CNN Money*. http://money.cnn.com/2011/11/09/smallbusiness/kiva_robots/, November 9, 2011.

[3] L. Ardissono, G. Petrone, G. Torta, and M. Segnan. M.: Mixed-initiative scheduling of tasks in user collaboration. In *In: Proc. of WEBIST 2012 - Eight International Conference on Web Information Systems and Technologies*, pages 342–351, 2012.

[4] Michael J. Barnes, Jessie Y.C. Chen, Florian Jentsch, and Elizabeth S. Redden. Designing effective soldier-robot teams in complex environments: training, interfaces, and individual differences. In *Proceedinggs of the International Conference on Engineering Psychology and Cognitive Ergonomics (EPCE)*, pages 484–493. Springer, 2011.

[5] Dimitris Bertsimas and Robert Weismantel. *Optimization over Integers*. Dynamic Ideas, Belmont, 2005.

[6] Elizabeth Blickensderfer, Janis A. Cannon-Bowers, and Eduardo Salas. Cross-training and team performance. making decisions under stress: Implications for individual and team training. *American Psychological Association*, pages 299–311, 1998.

[7] Jennifer Casper and Robin Roberson Murphy. Human-robot interaction in rescue robotics. *IEEE Transaction on Systems, Man, and Cybernetics (SMCS)*, 34(2):138–153, 2004.

[8] Jessie Y.C. Chen, Michael J. Barnes, and Zhihua Qu. Roboleader: an agent for supervisory control of mobile robots. In *Proceedings of the International Conference on Human-Robot Interaction (HRI)*, 2010.

[9] Andrew S Clare, Mary (Missy) L. Cummings, Jonathan P. How, Andrew K. Whitten, and Olivier Toupet. Operator objective function guidance for a real-time unmanned vehicle scheduling algorithm. 9:161–173, 2012.

[10] Mary (Missy) L. Cummings, Amy S. Brzezinski, and John D. Lee. Operator performance and intelligent aiding in unmanned aerial vehicle scheduling. *IEEE Intelligent Systems*, 22(2):52–59, March 2007.

[11] Edmund H. Durfee, James C. Boerkoel Jr., and Jason Sleight. Using hybrid scheduling for the semi-autonomous formation of expert teams. *Future Generation Computer Systems*, July 2013.

[12] Elliot E. Entin and Daniel Serfaty. Adaptive team coordination. *Human Factors*, 41:312–325, 1999.

[13] Dieter Fox. Adapting the sample size in particle filters through kld-sampling. *International Journal of Robotics Research (IJRR)*, 22:985–1003, 2003.

[14] Matthew C. Gombolay, Ronald J. Wilcox, and Julie A. Shah. Fast scheduling of multi-robot teams with temporospatial constrints. In *Proceedings of the Robots: Science and Systems (RSS)*, Berlin, Germany, June 24-28, 2013.

[15] Michael A. Goodrich, Bryan S. Morse, Cameron Engh, Joseph L. Cooper, and Julie A. Adams. Towards using UAVs in wilderness search and rescue: Lessons from field trials. *Interaction Studies, Special Issue on Robots in the Wild: Exploring Human-Robot Interaction in Naturalistic Environments*, 10(3):453–478, 2009.

[16] Masahiro Hamasaki, Hideaki Takeda, Ikki Ohmukai, and Ryutaro Ichise. Scheduling support system for academic conferences based on interpersonal networks. In *Proceedings of ACM Hypertext*, 2004.

[17] Thomas Haynes, Sandip Sen, Neeraj Arora, and Rajani Nadella. An automated meeting scheduling system that utilizes user preferences. In *Proceedings of the First International Conference on Autonomous Agents*, AGENTS '97, pages 308–315, New York, NY, USA, 1997. ACM.

[18] Guy Hoffman. Evaluating fluency in human-robot collaboration. In *International Conference on Human-Robot Interaction (HRI), Workshop on Human Robot Collaboration*, 2013.

[19] Guy Hoffman and Cynthia Breazeal. Effects of anticipatory action on human-robot teamwork: Efficiency, fluency, and perception of team. In *Proceedings of the International Conference on Human-Robot Interaction (HRI)*, pages 1–8, 2007.

[20] Eli R. Hooten, Sean T.Hayes, and Julie A. Adams. A comparison of communicative modes for map-based tasking. In *IEEE Internation Conference on Systems, Man, and Cybernetics*, 2011.

[21] Henry L. Jones, Stephen M. Rock, Dennis Burns, and Steve Morris. Autonomous robots in SWAT applications: Research, design, and operations challenges. *AUVSI*, 2002.

[22] Santiago Macho, Marc Torrens, and Boi Faltings. A multi-agent recommender system for planning meetings. In *Proceedings of Workshop on Agent-based Recommender Systems, Autonomous Agents 2000, ACM*, 2000.

[23] Colin F. Mackenzie, Yan Xiao, and Richard Horst. Video task analysis in high performance teams. *Cognition, Technology, and Work*, 6:139–147, 2004.

[24] Nicola Muscettola, Paul Morris, and Ioannis Tsamardinos. Reformulating temporal plans for efficient execution. In *Proceedings of the 6th International Conference on Principles of Knowledge Representation and Reasoning (KR&R)*, Trento, Italy, June 2-5, 1998.

[25] Stefanos Nikolaidis and Julie Shah. Human-robot cross-training: computational formulation, modeling and evaluation of a human team training strategy. In *Proceedings of the International Conference on Human-Robot Interaction (HRI)*, pages 33–40, 2013.

[26] Raja Parasuraman and Dietrich H. Manzey. Complacency and bias in human use of automation: An attentional integration. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 52:381–410, 2010.

[27] Raja Parasuraman, Mouloua Mustapha, and Brian Hilburn. Adaptive aiding and adpative task allocation enchance human-machine systems. *Automation technology and human performance: Current research and trends*, pages 119–123, 1999.

[28] Jason C. Ryan, Ashis Gopal Banerjee, Mary L. Cummings, and Nicholas Roy. Comparing the performance of expert user heuristics and an integer linear program in aircraft carrier deck operations. *IEEE Transaction on Cybernetics*, PP(9), August.

[29] Eduardo Salas, Jennifer E. Fowlkes, Renee J. Stout, Dana M. Milanovich, and Carolyn Prince. Does CRM training improve teamwork skills in the cockpit?: Two evaluation studies. *Human Factors*, 41:326–343, 1999.

[30] Julie Shah, James Wiken, Brian Williams, and Cynthia Breazeal. Improved human-robot team performance using chaski, a human-inspired plan execution system. In *Proceedings of the International Conference on Human-Robot Interaction (HRI)*, pages 29–36, 2011.

[31] C. Volpe, J. Cannon-Bowers, E. Salas, and P. Spector. The impact of cross training on team functioning. *Human Factors*, 38:87–100, 1996.

[32] Haoqi Zhang, Edith Law, Rob Miller, Krzysztof Gajos, David Parkes, and Eric Horvitz. Human computation tasks with global constraints. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 217–226, New York, NY, USA, 2012. ACM.