

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) 06-08-2015		2. REPORT TYPE Conference Proceeding		3. DATES COVERED (From - To) -	
4. TITLE AND SUBTITLE A Human Factors Analysis of Proactive Support in Human-robot Teaming			5a. CONTRACT NUMBER W911NF-13-1-0023		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER 611102		
6. AUTHORS Yu Zhang, Vignesh Narayanan, Tathagata Chakraborti, Subbarao Kambhampati			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES Arizona State University P.O. Box 876011  Tempe, AZ 85287 -6011			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 62486-NS.16		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT It has long been assumed that for effective humanrobot teaming, it is desirable for assistive robots to infer the goals and intents of the humans, and take proactive actions to help them achieve their goals. However, there has not been any systematic evaluation of the accuracy of this claim. <del>On the face of it, there are several ways a proactive robot</del>					
15. SUBJECT TERMS Human-factors, human-robot teaming					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Subbarao Kambhampati
UU	UU	UU	UU		19b. TELEPHONE NUMBER 480-965-0113

## Report Title

A Human Factors Analysis of Proactive Support in Human-robot Teaming

### ABSTRACT

It has long been assumed that for effective human-robot teaming, it is desirable for assistive robots to infer the goals and intents of the humans, and take proactive actions to help them achieve their goals. However, there has not been any systematic evaluation of the accuracy of this claim. On the face of it, there are several ways a proactive robot assistant can in fact reduce the effectiveness of teaming. For example, it can increase the cognitive load of the human teammate by performing actions that are unanticipated by the human. In such cases, even though the teaming performance could be improved, it is unclear whether humans are willing to adapt to robot actions or are able to adapt in a timely manner. Furthermore, misinterpretations and delays in goal and intent recognition due to partial observations and limited communication can also reduce the performance. In this paper, our aim is to perform an analysis of human factors on the effectiveness of such proactive support in human-robot teaming. We perform our evaluation in a simulated Urban Search and Rescue (USAR) task, in which the efficacy of teaming is not only dependent on individual performance but also on teammates' interactions with each other. In this task, the human teammate is remotely controlling a robot while working with an intelligent robot teammate 'Mary'. Our main result shows that the subjects generally preferred Mary with the ability to provide proactive support (compared to Mary without this ability). Our results also show that human cognitive load was increased with a proactive assistant (albeit not significantly) even though the subjects appeared to interact with it less.

**Conference Name:** IEEE/RSJ Intl. Conference on Intelligent Robots and Systems

**Conference Date:** September 28, 2015

# A Human Factors Analysis of Proactive Support in Human-robot Teaming

Yu Zhang, Vignesh Narayanan, Tathagata Chakraborti and Subbarao Kambhampati

**Abstract**—It has long been assumed that for effective human-robot teaming, it is desirable for assistive robots to infer the goals and intents of the humans, and take proactive actions to help them achieve their goals. However, there has not been any systematic evaluation of the accuracy of this claim. On the face of it, there are several ways a proactive robot assistant can in fact reduce the effectiveness of teaming. For example, it can increase the cognitive load of the human teammate by performing actions that are unanticipated by the human. In such cases, even though the teaming performance could be improved, it is unclear whether humans are willing to adapt to robot actions or are able to adapt in a timely manner. Furthermore, misinterpretations and delays in goal and intent recognition due to partial observations and limited communication can also reduce the performance. In this paper, our aim is to perform an analysis of human factors on the effectiveness of such proactive support in human-robot teaming. We perform our evaluation in a simulated Urban Search and Rescue (USAR) task, in which the efficacy of teaming is not only dependent on individual performance but also on teammates’ interactions with each other. In this task, the human teammate is remotely controlling a robot while working with an intelligent robot teammate ‘Mary’. Our main result shows that the subjects generally preferred Mary with the ability to provide proactive support (compared to Mary without this ability). Our results also show that human cognitive load was increased with a proactive assistant (albeit not significantly) even though the subjects appeared to interact with it less.

## I. INTRODUCTION

The efficacy of teaming [8] is not only dependent on individual performance, but also on teammates’ interactions with each other. It has long been assumed that for effective human-robot teaming, it is desirable for assistive robots to infer the goals and intents of the humans, and take proactive actions to help them achieve their goals. For example, the ability of goal and intent recognition is considered to be required for an assistive robot to be socially acceptable [22], [5], [16], [2], [24]. This claim is also assumed in other human-robot teaming tasks, such as collaborative manufacturing [25] and urban search and rescue (USAR) [23]. However, there has not been any systematic evaluation of the accuracy of this claim.<sup>1</sup>

<sup>\*</sup>This work was supported in part by the ARO grant W911NF-13-1-0023, and the ONR grants N00014-13-1-0176, N00014-13-1-0519 and N00014-15-1-2027.

The authors would like to thank Nathaniel Mendoza for help with the simulator as well as the anonymous participants in the study.

The authors are with the Department of Computer Science and Engineering, Arizona State University, Tempe, AZ 85281, USA {yzhan442, vnaray15, tchakra2, rao}@asu.edu

<sup>1</sup>The authors in [13] considered anticipatory action in interaction scenarios involving repetitive actions and the task settings are for human-robot teaming with proximal interactions.

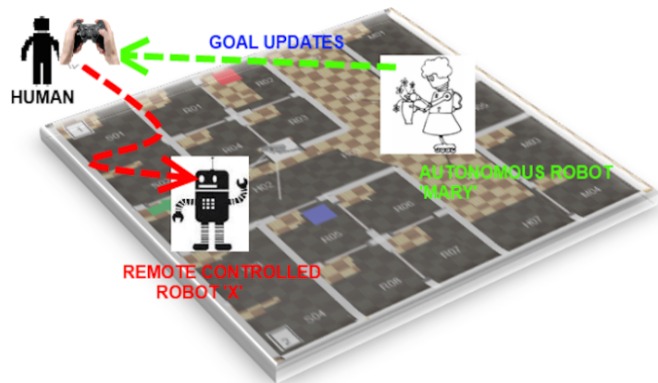


Fig. 1. Illustration of our USAR task in which the human teammate remotely controls a robot while working with an intelligent robot ‘Mary’. We intend to compare Mary with and without a proactive support ability.

There are several ways a proactive robot assistant can in fact reduce the effectiveness of teaming. For example, it can increase the cognitive load of the human teammate by performing actions that are unanticipated by the human. In such cases, even though the teaming performance could be improved, it is unclear whether humans are willing to adapt to robot actions or are able to adapt in a timely manner. Furthermore, misinterpretations and delays in goal and intent recognition due to partial observations and limited communication can also reduce performance. For example, consider a case in which you want to make an omelet and need eggs to be fetched from the fridge. Even if an assistive robot has started to fetch the eggs (after recognizing your intent), you may decide that the robot is too slow and fetch the eggs by yourself although you could improve the performance by letting the robot fetch the eggs while you preheat the pan. On the other hand, adapting to the robot’s actions in such scenarios to improve teaming performance can increase the human cognitive load, which leads to unsatisfactory teaming experience. These conflicting factors make us investigate the utility of Proactive Support (PS) in human-robot teaming.

In this paper, we start this investigation in a simulated USAR task with a general way to implement the proactive support ability on a robot in similar scenarios. Previous work [13] that investigates the effects of this ability is restricted to human-robot teaming with more proximal interactions. Meanwhile, to maintain the generality of this task, we only introduced a few necessary simplifications. In our task, the human teammate is remotely controlling a robot while working with an intelligent robot ‘Mary’ (as shown in Fig. 1). The human-robot team is deployed during the early phase

of an emergency response where they need to search, rescue and provide preliminary treatment to casualties.

This USAR scenario considers many of the complexities (e.g., partial observations) that often occur in real-world USAR tasks and we intend to learn whether these complexities influence the overall evaluation of the intelligent robot (i.e., Mary) with a proactive support (PS) ability, compared to Mary without this ability. We also aim to investigate the various trade-offs, e.g., mental workload and situation awareness through this human factors study.

## II. RELATED WORK

There are many early works on goal and intent recognition (e.g., [15], [14], [4]). More recently, a technique to compile the problem of plan recognition into a classical planning problem is provided [20]. There is also a rich literature in the area of plan adaptation, which handles how robots plan under human-introduced constraints (e.g., social rules [24]). Using simple temporal networks (STNs), there has been development in efficient dispatchers that perform fast, online, least-commitment scheduling adaptation [6]. There are also a number of adaptation techniques that focus on integrated planning and execution [7], [21], [1].

There are existing systems that combine both goal and plan recognition and plan adaptation to achieve a proactive support ability on robots. In [13], [12], the authors propose a cost based anticipatory adaptive action selection mechanism for a robotic teammate to make decisions based on the confidence of the action's validity and relative risk. However, only repetitive tasks are considered and the task settings are for human-robot teaming with more proximal interactions compared to that in USAR scenarios. In [5], a human-aware planning paradigm is introduced where the robot only passively interacts with the human by avoiding conflicts with the recognized human plan. In USAR scenarios, it is also desirable for the robot to proactively provide support to the human. A recent paper proposes a planning for serendipity paradigm in which the authors investigate planning for stigmergic collaboration without explicit commitments [3]. In [17], the authors propose a unified approach to concurrent plan recognition and execution for human-robot teams, in which they represent alternative plans for both the human and robot, thus allowing recognition and adaptation to be performed concurrently and holistically. However, the limitation is that the plan choices must be specified a priori instead of dynamically constructed based on the current goal and intent of the human. This renders the approach impractical for real-world scenarios since even moderate number of choices (i.e., branching factors) can make the approach infeasible.

Part of our goal is to provide a general way to achieve a proactive support ability in scenarios that are similar to our USAR task, in which the task is composed of subtasks with priorities that are dependent on the current situation. Note that a framework to achieve general proactive support can be arbitrarily complex depending on the task and level of support that is needed. In our work, similar to [23], we use the plan recognition technique in [20] and then feed its

outputs to a planner which determines the priorities of the subtasks and computes a plan accordingly. The main goal of this work is to start the investigation of human factors for proactive support in various human-robot teaming scenarios.

Regarding the benefits of automation in human-robot teaming, it is well known that automation can have both positive and negative effects on human performance. Empirical proofs have been provided in four main areas: mental workload, situation awareness, complacency and skill degradation [19]. We also aim to study the influence of proactive support on these factors in our USAR task.

## III. BACKGROUND

### A. USAR Task Settings Overview

In our simulated USAR task, the human and intelligent robot (i.e., Mary) share the same set of candidate goals (i.e., subtasks), and the overall team goal is to achieve them all (which will be distributed among the human and Mary). These goals are not independent of each other. In particular, the priorities of goals are dependent on which goals are achieved in the current situation. Given these task settings, we aim to investigate the influence of a proactive support (PS) ability on a robot. We compare two cases: Mary (i.e., the intelligent robot) has a PS ability and Mary does not have this ability. During the task execution, in both cases, Mary chooses her own goal to maximize the teaming performance accordingly to the human's current goal. When Mary does not have a proactive support ability, she can only know the human's current goal when the human explicitly communicates it to her. When Mary has this ability, if the human does not inform Mary of his/her current goal,<sup>2</sup> Mary can infer it based on her observations. To summarize, Mary in both cases can adapt to human goals while Mary with a PS ability can adapt in a more "proactive" fashion (hence proactive support). Finally, in both cases, Mary has an automated planner (see a brief description below) that can create a plan to achieve her current goal and she can autonomously execute the plan.

### B. Automated Planner

In our settings, a task or subtask is compiled into a problem instance for an automated planner to solve. The planner creates a plan by connecting an initial state to a goal state using agent actions. A planning problem can be specified using a planning domain definition language (PDDL) [11]. Depending on the task, there are many extensions of PDDL (e.g., [9], [10]) that can incorporate various modeling requirements. We use the extension of PDDL described in [9] to model the USAR domain. Using an automated planner allows an agent to reason directly about the goal. Human factors study on the incorporation of automated planners for human-robot teaming has appeared previously in [18].

<sup>2</sup>In both cases, when the human (optionally) informs Mary of his/her current goal, it is used directly by Mary assuming that this information is accurate.

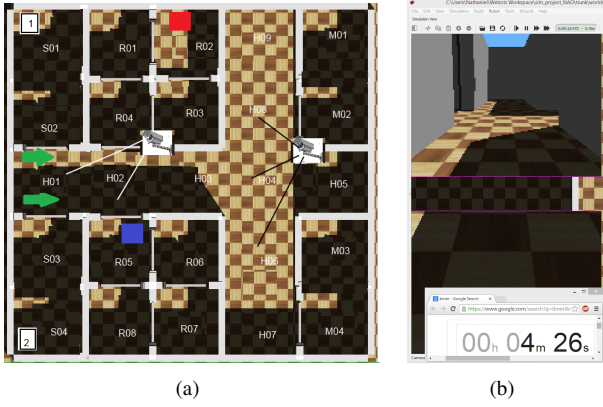


Fig. 2. (a) Simulated Environment for our USAR task. (b) The environment (from robot X's cameras) that the human subject actually sees.

### C. Goal and Intent Recognition

To recognize the human intents and goals, assuming that humans are rational, we use the technique in [20]. In our task, Mary maintains a belief of the human's current goal (denoted by  $G_X$ ) as a *hypothesis goal set*  $\Psi_X$ , in which  $\Psi_X$  corresponds to all remaining candidate goals. Given a sequence of observations  $\theta$  that are obtained periodically from sensors (on Mary or fixed in the environment), the probability distribution  $\Theta$  over  $G \in \Psi_X$  is recomputed using a Bayesian update  $P(G|\theta) \propto P(\theta|G)$ , where the prior is approximated by the function  $P(\theta|G) = 1/(1 + e^{-\beta\Delta(G,\theta)})$  in which  $\Delta(G, \theta) = C_p(G-\theta) - C_p(G+\theta)$ .  $C_p(G+\theta)$  and  $C_p(G-\theta)$  represent the cost of the optimal plan to achieve  $G$  with and without the observation of  $\theta$ , respectively. Having known the probability distribution  $\Theta$ , the goal that has the highest probability is assumed to be the current goal of the human. This goal is correspondingly taken out of the consideration of Mary and Mary then adapts her current goal if necessary (from her remaining goals) to optimize the teaming performance. Mary then makes a plan using an automated planner described previously to achieve her current goal.

## IV. STUDY DESIGN

### A. Hypotheses

We aim to investigate the following hypotheses:

- *H1*) Mary with a proactive support (PS) ability enables more effective teaming (e.g., less communication and more efficiency) in our task settings.
- *H2*) Mary with a PS ability increases human mental workload (e.g., due to unanticipated actions from Mary).

In our study, we also make efforts to maintain the task settings as general as possible. For a discussion on the generalization of the results, refer to the conclusion section.

### B. Environment

Fig. 2(a) shows the simulated environment (created in Webots) in our USAR task, which represents the floor plan of an office building where a disaster occurs (e.g., a fire). Fig. 2(b) is the visual feedback from the remotely controlled robot

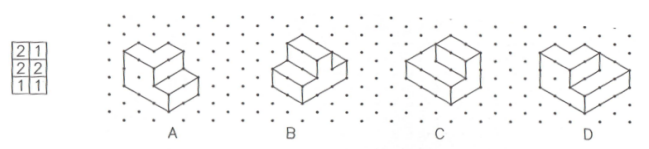


Fig. 3. Example puzzle problem used in our USAR task.

(i.e., robot X in Fig. 1) that the human subject actually sees. The environment is organized as segments, and each segment is identified by a unique label (e.g., R01). Furthermore, the segments are grouped into four regions: medical kit storage region (represented by segments starting with 'S'), casualty search region (starting with 'R'), medical room region where treatment (or triage) is performed (starting with 'M'), and the hallway region (starting with 'H'). Each region can be accessed via a door that connects to a hallway segment and R regions are further divided into rooms that are also connected by doors. The doors are initially closed and can be pushed open by the robots. The doors remain open after being pushed open. Both the remotely-controlled robot (denoted by 'X') and Mary work inside this environment. There are two networked CCTV cameras that Mary can obtain observations from and the field of views of these cameras are also shown in Fig. 2(a).

### C. Task Settings

The overall team goal is to find and treat all the casualties in the environment, which includes searching for casualties in the R regions, carrying casualties to medical rooms, fetching medical kits and performing triages. In Fig. 2(a), the two colored boxes (i.e., red and blue) in R regions represent casualties and the white boxes in S regions represent medical kits.

We impose two constraints on the agents: 1) either robot X or Mary can carry only one medical kit or one casualty at one time. 2) The triage can only be performed by robot X for which the human subject needs to solve a few puzzle problems (see Fig. 3 for an example) in 2 minutes. Out of the two casualties, we assume that one is critically injured (i.e., the red box in R02) who should be treated immediately after being found. The other one is lightly injured (i.e., the blue box in R05). It is also assumed that a medical room can only accommodate one casualty and each medical kit can only be used towards one casualty.

### D. Interface Design

In this USAR task, the human subject needs to manually control robot X while interacting with Mary. To create a more realistic USAR environment, the human subject only has access to the visual feeds from robot X. In other words, the human subject can only observe the part of the environment from robot X's "eyes" (i.e., two cameras, one mounted above the other).

The interaction interface between the human subject and robot X is shown in Fig. 4. More specifically, robot X displays a list of applicable actions that it can perform given



```

.....
1) pushdoor rob h02 r04
2) pushdoor rob h02 r05
3) move rob h02 h01
4) move rob h02 h03
.....
Select one action (or) Press 'i' to inform Mary about your current goal
i
.....
1) going to fetch medkit 1 from the top storage room
2) going to fetch medkit 2 from the bottom storage room
3) searching for casualty in top area
4) searching for casualty in bottom area
5) bringing medkit 1 to top medical room
6) bringing medkit 2 to top medical room
7) bringing medkit 1 to bottom medical room
8) bringing medkit 2 to bottom medical room
9) bringing highly critical casualty (red box) to top medical room
10) bringing less critical casualty (blue box) to top medical room
11) bringing highly critical casualty (red box) to bottom medical room i
12) bringing less critical casualty (blue box) to bottom medical room i
.....
Select your current goal (or) Press 'r' to return
r
4
Goal selected :4
.....
1) pushdoor rob h02 r04
2) pushdoor rob h02 r05
3) move rob h02 h01
4) move rob h02 h03
.....
Select one action (or) Press 'i' to inform Mary about your current goal

```

Fig. 4. Interaction interface between the human subject and robot X.

```

Mary's current goal:
.....
bringing medkit 2 to top medical room
.....
Mary's current goal interfere with human controller's (your) current goal:
How do you want Mary to execute ?
.....
1) Make Mary continue executing her current goal
2) Make Mary drop medkit/ casualty in the current segment
3) Make Mary drop medkit/casualty in the actual segment where it was fetched from
.....
2
Your selection :2
.....
Mary is dropping medkit in segment h02
.....
Mary's current goal:
.....
searching for casualty in top area
.....

```

Fig. 5. Interaction interface between the human subject and Mary.

the current state. The human subject interacts with robot X to choose an action from the list of applicable actions. When the chosen action is completed by X, the interaction interface displays the next set of actions. This process is repeated until the task is finished (i.e., all the casualties are found and treated). Following are the list of all possible action types that the human can choose. Compare the list with that shown in Fig 4. This interface also allows the human subject to optionally inform Mary about his/her current goal so that Mary can remove it from consideration and adapt her goal accordingly when necessary.

- move X H01 H02 - Move robot X from hallway segment H01 to hallway segment H02.
- pushdoor X R01 R02 - Push the door between room R01 and room R02.
- grab\_medkit X S01 - Grab the medical kit from storage room S01.
- carry\_casualty X R01 - Carry the casualty at room R05.
- drop\_medkit X M01 - Drop the medical kit in medical room M01.
- lay\_down\_casualty X M01 - Lay down the casualty in medical room M01.
- perform\_triage X M01 - Perform medical triage in medical room M01.
- Press 'i' - Inform Mary about the human subject's current or intended goal. (A list of all remaining candidate goals will be displayed to be chosen.)

Note that these actions are modeled to respect the constraints that we discussed in Sec. IV-C. For example, lay\_down\_casualty X M01 is only available when there is no other casualties in medical room M01; perform\_triage X M01 is only available when there is a casualty and a medical kit in M01.

The interaction interface between the human subject and Mary is shown in Fig. 5. This interface is first used by Mary to update the human subject about her current goal. When the human subject wants to take over the goal that Mary is



Fig. 6. Experimental setup in the USAR task

acting to achieve, this interface is also used to display the choices (to be selected by the human subject) for Mary to terminate her current (uncompleted) goal.

### E. Study Setup and Flow

The study was set up in our lab space, similar to that shown in Fig. 6. Before the beginning of the task, the human subject is given the floor plan without the annotations of the casualties (i.e., colored boxes). Furthermore, the human subject is informed that there are two casualties (that cannot move) and they are located inside the casualty search regions. However, no information about their exact locations is provided (i.e., which rooms the casualties are in). The human subject is also informed that the casualty that is represented by a red box is seriously injured, and should be treated as soon as possible. Note that Mary has no more information than the human subject. The remotely controlled robot X and Mary start in the same segment H01, which is specified by the green arrows.

Subjects were assigned alternately to team up with either Mary with a PS ability or without. Each subject is only

allowed to take part in one experimental trial to avoid performance fluctuation due to experience. All subjects completed the consent form before participating in the study. Prior to each run, the subject was asked to read the instruction materials that contain the background knowledge and the above information. The subject was then exposed to the simulator and the interface and was asked to experiment with them to gain some familiarity. The subject was asked to collaborate with Mary to find and treat the two casualties. After the trial, the subject was asked to complete a questionnaire (in Likert scale).

#### F. Example Scenario

Next, we walk through an example scenario in our USAR task. Consider a scenario in which the human subject found the critically injured casualty and the current goal ( $\mathbb{G}_X$ ) of the human subject becomes ‘bring the critically injured casualty to the top medical room in Fig. 2(a):

```
goal(X, 'bring the critically injured
casualty to the top medical room') =
{ (at critically_injured_casualty M01)}
```

However, assume that the human subject failed to inform Mary of his/her current goal. Also, assume the following states for the medical kits:  $\{(at\ med\_kit\_1\ S01), (at\ med\_kit\_2\ S04)\}$ , and that Mary at that time is still searching the casualties in the other casualty search region. When robot  $X$  enters the field of view of the CCTV cameras the action and state of  $X$  are detected by the cameras and are fed to Mary as observations. In this example, some of robot  $X$ 's actions, such as  $\{(move\ X\ H02\ H03), (move\ X\ H04\ H08)\}$  will be observed by Mary, which triggers the goal and intent recognition process. After computing the probability distribution  $\Theta$  for all goals in the candidate goal set for the human, the goal that has the higher probability (and falls above a pre-specified threshold) is assumed to be the current goal of the human ( $\mathbb{G}_X$ ), which in this case is ‘bring the critically injured casualty to the top medical room’. Mary now knows that the critically injured casualty has been found and can remove this goal from her own candidate goal set.

Furthermore, given this information, Mary recomputes the priorities of the remaining goals in the current situation and adapts her goal accordingly. In particular, although the searching task is still undergoing, Mary realizes that in this case helping the human subject by bringing a medical kit to  $M01$  would achieve a better utility for the team. Note that should the casualty found by the human subject be lightly injured instead, Mary would decide to continue her search; also, should the casualty found by the human subject be lightly injured but the critically injured casualty has already been treated, Mary would choose to help the human fetch the medical kit. Note also that in the case that Mary does not have a PS ability, the above update can only occur in a timely manner if the human subject chooses to inform Mary about his/her current goal. In our running example, the goal

that Mary chooses is:

```
goal( $\mathbb{G}_M$ , 'bring med_kit_1 to the top
medical room') =
{(at med_kit_1 M01)}
```

Having chosen her current goal  $\mathbb{G}_M$ , Mary then uses an automated planner to generate a plan ( $\Pi_M$ ) that achieves the goal. Meanwhile, Mary will update the human subject with her current goal. Assuming that Mary is at segment  $H01$  at the time, the following plan would be generated:

```
 $\Pi_M = \langle (pushdoor\ Mary\ H01\ S03),
(move\ Mary\ H01\ S03),
(move\ Mary\ S03\ S04),
(grab\_medkit\ Mary\ S04),
(move\ Mary\ S04\ S03),
(move\ Mary\ S03\ H01),
(move\ Mary\ H01\ H02),
(move\ Mary\ H02\ H03),
(move\ Mary\ H03\ H04),
(move\ Mary\ H04\ H08),
(pushdoor\ Mary\ H08\ M02),
(move\ Mary\ H08\ M02),
(move\ Mary\ M02\ M01),
(drop\_medkit\ Mary\ M01)\rangle$ 
```

Note that various other scenarios can arise in this task, which may not always favor Mary with a PS ability. For example, the human subject may decide to deliver the medical kits to the medical rooms even before finding any casualties. or the human subject may walk robot  $X$  to the medical room empty-handed. These can confuse the goal and intent recognition process on Mary and lead to reduced teaming performance. Although not all of these scenarios occurred during our experimental study, they demonstrate the conflicting factors for proactive support in human-robot teaming tasks. It is also clear that these tradeoffs are dependent on the task and robot settings, which require more investigations in future work.

## V. RESULTS

The study was performed over 4 weeks and involved 16 volunteers (9 males, 7 females), Volunteers have ages with  $M = 24$  and  $SD = 1.15$ . Subjects were recruited from students on campus. Due to the requirement of understanding English instructions, subjects must indicate that they are confident with English communication skills before taking part in the study. We also asked about the subject's familiarity with computers ( $M = 6.56$ ,  $SD = 0.63$ ), robots ( $M = 4.19$ ,  $SD = 0.91$ ), puzzle problems ( $M = 3.19$ ,  $SD = 0.83$ ) and computer gaming ( $M = 4.69$ ,  $SD = 1.49$ ), in seven-point scales after the study (with 1 being least familiar and 7 being most familiar). The subjects reported familiarity with computers, but not so much with robots, puzzle problems or computer gaming.

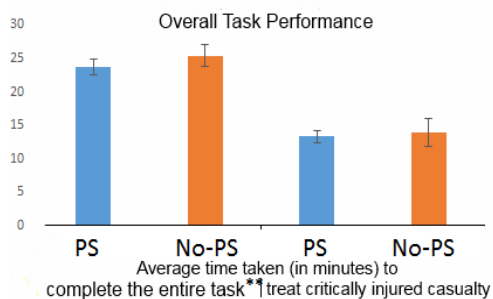


Fig. 7. Results for objective performance and measures. \* denotes  $p < 0.05$ , \*\* denotes  $p < 0.01$ , \*\*\* denotes  $p < 0.001$ .

### A. Measurement

A post-study questionnaire is used to evaluate three of four areas that are often used to assess automated systems: mental workload, situation awareness, and complacency [19]. Furthermore, we also use the questionnaire to evaluate several psychological distances between individuals and the environment (including robots), which include immediacy, effectiveness, likability and trust. Immediacy describes how realistic the subject felt about the task and Mary. Effectiveness describes the subject’s feeling about how effective the subject considered Mary as a teammate. Likability describes how likable the subject felt about Mary. Trust describes whether the subject felt that Mary was trustworthy. We also collect the subjects’ opinions on whether they considered that Mary should be improved (i.e., *improvability*).

One way fixed-effects ANOVA tests were performed to analyze the objective performance and measures, as well as the subjective questions. The fixed factor in the tests is the type of Mary, the intelligent robot, which is either Mary with a PS ability or without (denoted by No-PS).

### B. Objective Performance

We first investigate the objective performance and measures. The overall performance (presented in in Fig. 7) is evaluated based on the total time taken for the team to find and treat the critically injured casualty, and the total time taken for the team to finish the entire task (i.e., find and treat both casualties). It is interesting to observe that while there is a significant difference between PS and No-PS for the time taken to complete the entire USAR task ( $F(1, 14) = 8.34, p < 0.01$ ), we do not find any significant difference for treating the critically injured casualty. This may be due to the fact that humans are proficient at prioritizing goals. However, this may negatively impact the teaming performance since the subject may more often choose to neglect the help of Mary when he/she does not feel comfortable with entrusting Mary with important goals. This conjecture is also consistent with the results in Fig. 8, which is discussed next.

We provide a more detailed analysis of task performance in Fig. 8. We compare the average number of times the subject stopped Mary from executing her current goal and the average number of times the subject had goal conflicts with Mary. The results show that these numbers are generally

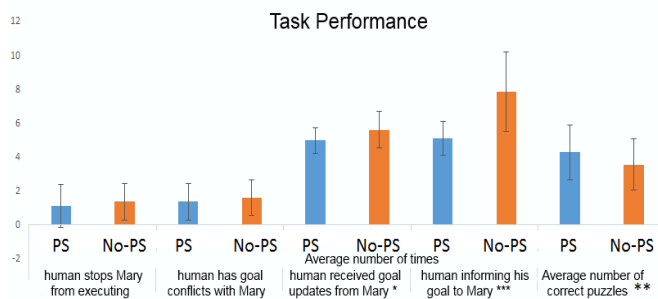


Fig. 8. Results for task performance and measures. \* denotes  $p < 0.05$ , \*\* denotes  $p < 0.01$ , \*\*\* denotes  $p < 0.001$ .

smaller for the PS case but we did not find any significant difference. However, we did find a significant difference for the average number of times the subject informed his/her goal to Mary ( $F(1, 14) = 18.27, p < 0.001$ ). This shows that the subject felt less necessity to inform Mary in the PS case. There is also a significant difference in the number of goal updates the subject received from Mary ( $F(1, 14) = 7.58, p < 0.05$ ), This confirms that Mary changed her goal less frequently in the PS case.

We also compare the accuracy of the puzzle problems for the triage operations. To discourage subjects from guessing the answers to the puzzle questions, they were told that each incorrect answer would give them negative scores. Our analysis, interestingly, shows a significant difference on this performance measure ( $F(1, 14) = 4.64, p < 0.01$ ), which suggests that the human mental workload may have been reduced in the PS case, which is not consistent with the second hypothesis (i.e.,  $H2$ ). Furthermore, as we show in the evaluation of subjective measures, this interpretation contradicts with the results there.

### C. Subjective Performance

In this section, we investigate the subjective performance based on the questionnaire (23 questions in total). For these 23 questions, we categorize them into 8 different (partially overlapping) groups. This includes 3 groups for evaluating automation: mental workload (3 items, Cronbach’s  $\alpha = 0.713$ ), situation awareness (1 item), and complacency (2 items, Cronbach’s  $\alpha = 0.769$ ). Furthermore, we also evaluate several psychological distances between the human subject and environment (including Mary), which include immediacy (1 item), effectiveness (7 items, Cronbach’s  $\alpha = 0.724$ ), likability (1 item), and trust (3 items, Cronbach’s  $\alpha = 0.871$ ). We also include improvability (1 item). The answers to the questions are in seven-point scales. The results are presented accumulatively in Fig. 9.

1) *Mental Workload*: For mental workload, we include questions that inquire about the ease of working with Mary, and questions to rate the subject’s mental workload to interact with Mary during the task. Although our analysis does not find any significant difference ( $p = 0.404$ ), the subjects still reported some difference in their mental workloads. This is an interesting result that confirms our hypothesis (i.e.,  $H2$ ):



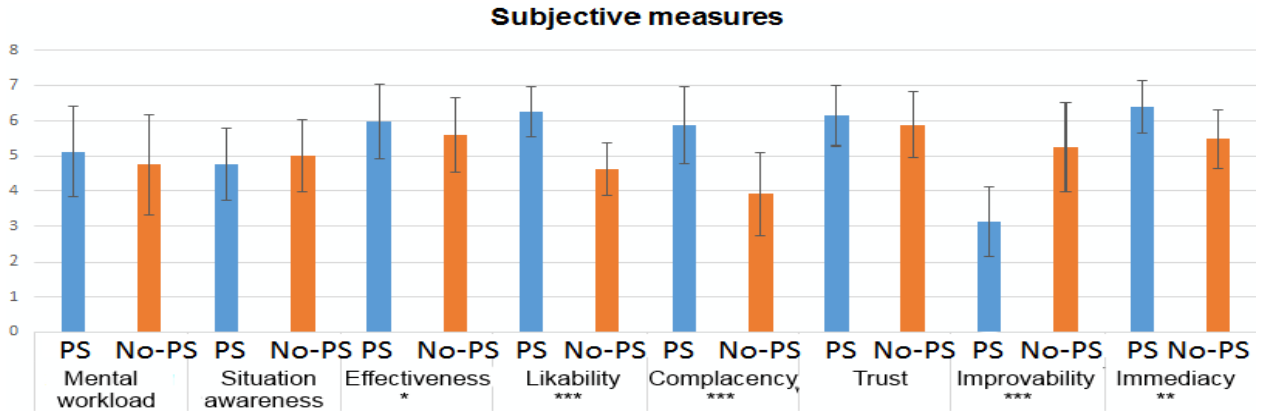


Fig. 9. Results for subjective measures. \* denotes  $p < 0.05$ , \*\* denotes  $p < 0.01$ , \*\*\* denotes  $p < 0.001$ .

although the PS ability enables more effective human-robot teaming, it also tends to increase the human mental workload at the same time. It is also worth noting that even though the subjects in the PS case reported increased mental workload, they also tended to perform well on the puzzle problems. This may be due to the fact that subjects felt less necessity to communicate with Mary and thus can concentrate more on these problems.

2) *Situation Awareness*: For situation awareness, we include questions that inquire about whether the subject felt that he/she had enough information to determine what the next goal should be. Our analysis does not show a significant difference ( $F(1, 14) = 2.78, p = 0.35$ ), although the subjects reported slightly more situation awareness in the No-PS case, which is consistent with the side effects of automation in general. Although the number of updates for the No-PS case was significantly more than that for the PS case, the fact that situation awareness of the subject was not reduced much in the PS case is encouraging. We attribute this to the fact that the subject still needed to occasionally interact with Mary when they had goals conflicts, and the subject could gain situation awareness through such interactions.

3) *Complacency*: For complacency, we include questions about the comfort and ease of the teaming, as well as how well the subject felt about their performance in the task. Our analysis shows a significant difference ( $F(1, 14) = 11.29, p < 0.001$ ). This is consistent with the objective performance and measures, which shows that the human subject generally felt more satisfied and confident working with Mary in the PS case. This is important for human-robot teaming.

4) *Immediacy, Effectiveness, Likability & Trust*: For immediacy, we include questions about how much the subject considered the simulated task as a realistic USAR task, and Mary as a teammate. Our analysis shows a significant difference ( $F(1, 14) = 11.63, p < 0.001$ ), which is consistent with our prior results.

For effectiveness, we include questions about the perceived effectiveness of the team, the balance of workload between the team members, and whether or not the subject felt that Mary performed expectedly. Our analysis shows a significant difference ( $F(1, 14) = 6.57, p < 0.05$ ). This result suggests

that the proactive support ability indeed increases teaming effectiveness.

For likability, we include questions about whether the subject felt that Mary was a good teammate. Our analysis shows a significant difference ( $F(1, 14) = 23.26, p < 0.001$ ), which suggests that the subjects preferred Mary with a PS ability for teaming.

For trust, we include questions about the evaluation of the Mary's trustworthiness with the assignments (or tasks) she took and with her updates during the task. Our analysis did not show any significant difference with  $F(1, 14) = 3.78, p = 0.072$ , although subjects in the PS case reported slightly higher trust.

5) *Improvability*: For improvability, we include questions about how much the subject felt that Mary could be improved, and how the subject evaluated his/her interaction with Mary. Our analysis shows a significant difference for improvability with  $F(1, 14) = 17.80, p < 0.001$ , which, again, suggests that the subjects preferred Mary with a PS ability.

#### D. Summary

In summary, our results are mostly consistent with our hypotheses. Our main result shows that the subjects generally preferred Mary with a PS ability. With the PS ability, the human cognitive load was indeed increased (albeit not significantly), even though the subjects appeared to interact less with Mary. More specifically, while the result on mental workload confirms our hypothesis, it also seems to be conflicting with the objective performance on the puzzle problems. This is likely due to the fact that the subject felt less necessity to interact with Mary in the PS case. Furthermore, given that situation awareness was not reduced significantly in the team with Mary having a PS ability, and that the subjects had positive feelings towards her, it seems to suggest that intelligent robots with a PS ability is welcomed in general. This is, of course, largely dependent on the fact that the subject's cognitive load is not increased significantly, which may change when the human needs to adapt to the robot's action more frequently in more complex tasks, and more communication may be needed. More investigations

are needed to be conducted in such scenarios where the task and robot settings largely differ.

## VI. CONCLUSIONS

In this paper, we aim to start the investigation of human factors for proactive support in human-robot teaming. We start in a simulated USAR task with a general way to implement the proactive support (PS) ability on a robot in similar scenarios in which the task is composed of subtasks with priorities that are dependent on the current situation. Meanwhile, to maintain the generality of this task, we only introduced a few necessary simplifications. However, given the richness of USAR scenarios, more in depth studies are required to generalize the conclusions to scenarios where the task and robot settings largely differ. In such cases, our plan recognition and plan adaptation approaches may also need to be extended to implement proactive support. Note that a framework to achieve general proactive support can be arbitrarily complex depending on the task and level of support that is needed (e.g., whether the support is active [13] or passive [5] and whether it is commitment sensitive or not [3]).

In our task, the human teammate is remotely controlling a robot while working with an intelligent robot Mary to search for and treat casualties. Our results show that, in general, the human teammates prefer to work with a robot that has a PS ability. However, our results also show that teaming with PS robots also increases the human's cognitive load, albeit not significantly. This is understandable since working with a proactive teammate may require more interactions and/or mental modeling on the human side in order to achieve better teaming performance. Furthermore, we also show that situation awareness when working with robots with a PS ability is not significantly reduced compared to working with robots without it. This seems to suggest that intelligent robots with a PS ability is welcomed in general.

## REFERENCES

- [1] Samir Alili, Matthieu Warnier, Muhammad Ali, and Rachid Alami. Planning and plan-execution for human-robot cooperative task achievement. *Proc. of the 19th ICAPS*, pages 1–6, 2009.
- [2] Filippo Cavallo, Raffaele Limosani, Alessandro Manzi, Manuele Bonaccorsi, Raffaele Esposito, Maurizio Di Rocco, Federico Pecora, Giancarlo Teti, Alessandro Saffiotti, and Paolo Dario. Development of a socially believable multi-robot solution from town to home. *Cognitive Computation*, 6(4):954–967, 2014.
- [3] Tathagata Chakraborti, Gordon Briggs, Kartik Talamadupula, Yu Zhang, Matthias Scheutz, David Smith, and Subbarao Kambhampati. Planning for serendipity. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2015.
- [4] Eugene Charniak and Robert P. Goldman. A bayesian model of plan recognition. *Artificial Intelligence*, 64(1):53 – 79, 1993.
- [5] M. Cirillo, L. Karlsson, and A. Saffiotti. Human-aware task planning for mobile robots. In *Advanced Robotics, 2009. ICAR 2009. International Conference on*, pages 1–7, June 2009.
- [6] Rina Dechter, Itay Meiri, and Judea Pearl. Temporal constraint networks. *Artificial intelligence*, 49(1):61–95, 1991.
- [7] Alberto Finzi, Félix Ingrand, and Nicola Muscettola. Model-based executive control through reactive planning for autonomous rovers. In *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, volume 1, pages 879–884. IEEE, 2004.
- [8] Terrence Fong, Illah Nourbakhsh and Clayton Kunz, Lorenzo Fluckiger, John Schreiner, Robert Ambrose, Robert Burrigge, Reid Simmons, Laura Hiatt, Alan Schultz, J. Gregory Trafton, Magda Bugajska, and Jean Scholtz. The peer-to-peer human-robot interaction project. *Space 2005*.
- [9] Maria Fox and Derek Long. Pddl2. 1: An extension to pddl for expressing temporal planning domains. *J. Artif. Intell. Res.(JAIR)*, 20:61–124, 2003.
- [10] Alfonso Gerevini and Derek Long. Plan constraints and preferences in pddl3. *The Language of the Fifth International Planning Competition. Tech. Rep. Technical Report, Department of Electronics for Automation, University of Brescia, Italy*, 75, 2005.
- [11] Malik Ghallab, Craig Knoblock, David Wilkins, Anthony Barrett, Dave Christianson, Marc Friedman, Chung Kwok, Keith Golden, Scott Penberthy, David E Smith, et al. Pddl-the planning domain definition language. 1998.
- [12] Guy Hoffman and Cynthia Breazeal. Cost-based anticipatory action selection for human-robot fluency. *Robotics, IEEE Transactions on*, 23(5):952–961, 2007.
- [13] Guy Hoffman and Cynthia Breazeal. Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction*, pages 1–8. ACM, 2007.
- [14] Henry A. Kautz. Reasoning about plans. chapter A Formal Theory of Plan Recognition and Its Implementation, pages 69–124. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1991.
- [15] Henry A. Kautz and James F. Allen. Generalized Plan Recognition. In *National Conference on Artificial Intelligence*, pages 32–37, 1986.
- [16] Uwe Köckemann, Federico Pecora, and Lars Karlsson. Grandpa hates robots - interaction constraints for planning in inhabited environments. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada.*, pages 2293–2299, 2014.
- [17] Steven James Levine and Brian Charles Williams. Concurrent plan recognition and execution for human-robot teams. In *Twenty-Fourth International Conference on Automated Planning and Scheduling*, 2014.
- [18] Vignesh Narayanan, Yu Zhang, Nathaniel Mendoza, and Subbarao Kambhampati. Automated planning for peer-to-peer teaming and its evaluation in remote human-robot interaction. In *ACM/IEEE International Conference on Human Robot Interaction (HRI)*, 2015.
- [19] Raja Parasuraman. Designing automation for human use: empirical studies and quantitative models. *Ergonomics*, 43(7):931–951, 2000.
- [20] Miquel Ramírez and Hector Geffner. Probabilistic plan recognition using off-the-shelf classical planners. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010, Atlanta, Georgia, USA, July 11-15, 2010*, 2010.
- [21] 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 6th international conference on Human-robot interaction*, pages 29–36. ACM, 2011.
- [22] E. A. Sisbot, L. F. Marin-Urias, R. Alami, and T. Simeon. A human aware mobile robot motion planner. *IEEE Transactions on Robotics*, 2007.
- [23] Kartik Talamadupula, Gordon Briggs, Tathagata Chakraborti, Matthias Scheutz, and Subbarao Kambhampati. Coordination in human-robot teams using mental modeling and plan recognition. In *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pages 2957–2962, Sept 2014.
- [24] Stevan Tomic, Federico Pecora, and Alessandro Saffiotti. Too cool for school - adding social constraints in human aware planning. In *Proceedings of the International Workshop on Cognitive Robotics (CogRob)*, 2014.
- [25] Vaibhav V Unhelkar, Ho Chit Siu, and Julie A Shah. Comparative performance of human and mobile robotic assistants in collaborative fetch-and-deliver tasks. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pages 82–89. ACM, 2014.