International Conference on Robotics and Biomimetics
December 5-10, 2014, Bali, Indonesia

Task Switching and Cognitively Compatible guidance for Control of Multiple Robots

Michael Lewis, *Member IEEE*, Shi-Yi Chien, Siddarth Mehrotra,
Nilanjan Chakraborty, *Member IEEE*, and Katia Sycara, *Fellow IEEE*

***Abstract*—Decision aiding sometimes fails not because following guidance would not improve performance but because humans have difficulty in following guidance as it is presented to them. This paper presents a new analysis of data from multi-robot control experiments in which guidance in a demonstrably superior robot selection strategy failed to produce improvement in performance. We had earlier suggested that the failure to benefit might be related to loss of volition in switching between robots being controlled. In this paper we present new data indicating that spatial, and hence cognitive proximity, of robots may play a role in making volitional switches more effective. Foraging tasks, such as search and rescue or reconnaissance, in which UVs are either relatively sparse and unlikely to interfere with one another or employ automated path planning, form a broad class of applications in which multiple robots can be controlled sequentially in a round-robin fashion. Such human-robot systems can be described as a queuing system in which the human acts as a server while robots presenting requests for service are the jobs. The possibility of improving system performance through well- known scheduling techniques is an immediate consequence. Two experiments investigating scheduling interventions are described. The first compared a system in which all anomalous robots were alarmed (Alarm), one in which alarms were presented singly in the order in which they arrived (FIFO) and a Control condition without alarms. The second experiment employed failures of varying difficulty supporting an optimal shortest job first (SJF) policy. SJF, FIFO, and Alarm conditions were compared. In both experiments performance in directed attention conditions was poorer than predicted. This paper presents new data comparing the spatial proximity in switches between robots selected by the operator (Alarm conditions) and those dictated by the system (FIFO and SJF conditions).**

I. INTRODUCTION

In the simplest case of multirobot control, an operator controls multiple independent robots interacting with each as needed. A foraging task [1] in which each robot searches its own region would be of this category. Control performance at such tasks can be characterized by the average demand of each robot on human attention [2]. Such operator interactions with a robot might be described as a sequence of control episodes in which an operator interacts with the robot for period of time (interaction time, IT) raising its performance above some upper threshold (UT) after which the robot is neglected for a period of time (neglect time, NT) until its performance deteriorates below a lower threshold (LT) when the operator must again interact with it. In practice the operator's task is even more complex. Humans are additionally included in robotic systems to perform tasks the automation cannot. The most common of these tasks is searching for targets in noisy displays such as remote video or aerial imagery.

Research in robot self-reflection [3] has progressed to the point that it is plausible to presume robots capable of reporting their own off normal conditions such as an inability to move or unsafe attitude. By focusing the operator's attention on robots needing interaction rather than requiring the operator to monitor for the failures, time spent monitoring can be eliminated increasing the number of robots that can be serviced over this interval. With robots informing the operator of their need for interaction the human-robot system becomes more like a queuing system in which the operator acts as the server and robot interaction requests as jobs. Using operations research methods the performance of such a queuing system might be further improved by prioritization of jobs or adjustment of service levels [4] to match current conditions. Deriving full benefit from such aiding, however, would require the ability to focus an operator's attention on a particular robot. We refer to the possibility that human attention might be closely directed in this manner without loss of cognitive efficiency as the *attention scheduling* hypothesis.

Alarms are commonly used in complex human-machine systems to direct human attention but usually in an open and unrestrictive way. Annunciator systems in nuclear power plants or aircraft cockpits typically alarm separately for each setpoint that has been exceeded allowing the human to prioritize and schedule attention among competing demands. Human-multirobot tasks exert similar competing demands on operators frequently requiring them to mix navigation, visual search, and status monitoring to accomplish their objectives. If operators can manage their own attentional resources to avoid damaging interruptions and/or exploit common situational elements among tasks these advantages might outweigh benefits available from externally directed attention.

Experiment I tests the *attention scheduling* hypothesis by comparing operators performing a multirobot foraging task without alarms for robot failures, with all alarms available (Alarm), or with a first-in-first-out (FIFO) queue

Manuscript received August 19, 2014. This work was supported in part by the U.S. Office of Naval Research Grant N0001409-10680

M. Lewis is with the University of Pittsburgh, Pittsburgh, PA 15260 e-mail: ml@sis.pitt.edu.

S.Chien is with the University of Pittsburgh, Pittsburgh, PA 15260 e-mail: gschien@gmail.com

S. Mehrotra was with Carnegie Mellon University, Pittsburgh, PA 15213. He is now with Citicorp, New York, NY email: siddharthmehrotra11@gmail.com

N. Chakraborty was with Carnegie Mellon University, Pittsburgh, PA 15213. He is now with SUNY Stony Brook, NY email: nilanjan.chakraborty@stonybrook.edu

K. Sycara is with Carnegie Mellon University, Pittsburgh, PA 15213. email: katia@cs.cmu.edu

978-1-4799-7397-2/14/$31.00 © 2014 IEEE

making only a single alarm available at a time. Effects were measured for both the primary task of searching for and identifying victims and the secondary task of identifying and restoring failed robots. Because all failures were of the same difficulty, the order in which they were serviced should make no difference so under the *attention scheduling* hypothesis the FIFO and Alarm conditions should produce equivalent performance.

Experiment II extends the test to a condition under which the attention scheduling hypothesis would predict superior performance for directed attention. The shortest job first (SJF) discipline is a provably optimal policy for maximizing throughput in a queuing system [5]. Using this policy to direct human attention, therefore, should lead to superior performance under the *attention scheduling* hypothesis providing the undirected operators did not follow precisely the same policy. This experiment compares SFJ, FIFO, and Alarm conditions with *attention scheduling* hypothesis predictions that SJF should produce the best performance followed by Alarm provided that operators did better than random (FIFO) in selecting robots to be serviced.

II. METHODS

A. USARSim and MrCS

The reported experiments were conducted using the USARSim robotic simulation with simulated Pioneer P3-AT robots performing an Urban Search and Rescue (USAR) foraging task. USARSim is a high-fidelity simulation of USAR robots and environments developed as a research tool for the study of human-robot interaction (HRI) and multi-robot coordination. USARSim supports HRI by accurately rendering user interface elements (particularly camera video), accurately representing robot automation and behavior, and accurately representing the remote environment that links the operator's awareness with the robot's behaviors. USARSim uses Epic Games' UnrealEngine3 to provide a high fidelity simulator at low cost and also serves as the basis for the Virtual Robots Competition of the RoboCup Rescue League. Other sensors including sonar and audio are also accurately modeled.

MrCS (Multi-robot Control System), a multi-robot communications and control infrastructure with accompanying user interface, developed for experiments in multirobot control and RoboCup competition [6] was used in these experiments. MrCS provides facilities for starting and controlling robots in the simulation, displaying multiple camera and laser output, and supporting inter-robot communication.

Figure 1 shows the MrCS user interface in the Alarm condition. Thumbnails of robot camera feeds are shown on the top, a video feed of interest in the bottom right. A GUI element in the middle right allows teleoperation and camera pan and tilt. Current locations and paths of the robots are shown on the Map Viewer (middle) which also allows operators to mark victims. The team status window (left) for the Alarm condition shows each robot's current status and briefly summarizes any problem. Green indicates the robot is in autonomous condition and functioning safely, yellow indicates an abnormal condition, such as stuck at a corner. When a robot is manually controlled, its tile turns white.

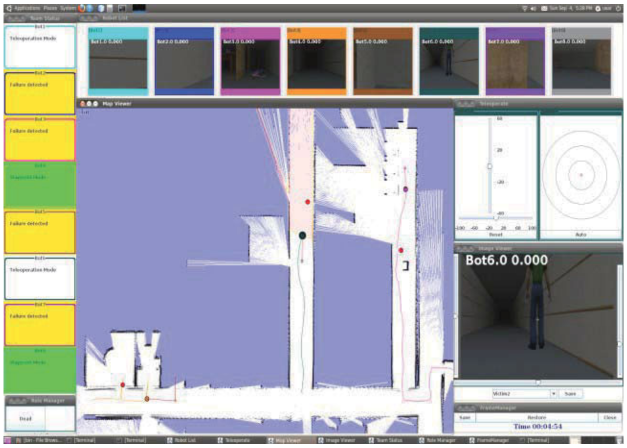

Fig. 1. MrCS Alarm condition with status bar on left

The operator selects the robot to be controlled from either the team status window or camera thumbnail. Figure 2 shows the team status window for the forced queue conditions (FIFO/SJF) in which robots in abnormal states are presented one at a time. Additional alarms can only be reviewed after the presenting problem is resolved. To avoid "clogging" the status window with an unrecoverable failure, operators have an alternative in the Dead button (bottom left). Once switched off, the robot will stop reporting and no longer be scheduled. The status panel is removed in the Control condition requiring operators to monitor the Map Viewer and thumbnails to identify malfunctioning robots.

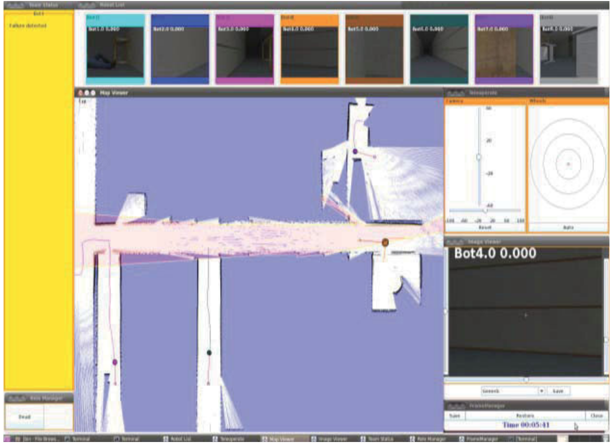

Fig. 2. Decision aid condition display (FIFO-/Priority-Queue).

B. USAR Foraging Task

When an operator detects a victim in a thumbnail, a complex sequence of actions is initiated. The operator first needs to identify the robot and select it to see the camera view in a larger window and to gain the ability to stop or teleoperate the robot. After the user has successfully selected a robot, it must be located on the map by matching the window border color or numerical label. Next the operator must determine the orientation of the robot and its camera using cues such as prior direction of motion and matching landmarks between camera and map views. To gain this information the operator may choose to teleoperate the selected robot to locate it on the map, determine its orientation through observing the direction of movement, or simply to get a better viewing angle. The operator must then estimate the location on the map corresponding to the victim in the camera view. If "another" victim is marked nearby, the operator must decide whether the victim she is preparing to mark has already been recorded on the map.

Detecting and restoring a failed robot follow a similar time course: identifying the failed robot on the map and selecting it, then teleoperating it to its next waypoint where the automation can resume control.

C. *Experimental Conditions*

The selected USAR environment was an office like hall with many rooms full of obstacles like chairs and desks. Victims were evenly distributed within the environment. Maps were rotated by 90º and each robot entered the environment from different locations on each trial. Because the laser map is built up slowly as the environment is explored and the office like environment provides few distinctive landmarks, there was little opportunity for participants to benefit from prior exposure to the environment. Robots followed predefined paths of waypoints, similar to paths generated by an autonomous path planner [7] to explore the map. All robots traveled paths of the same distance encountering the same number of victims and failures in each designed path. Upon reaching a failure point the operator needed to assume manual control to teleoperate the robot out of its predicament to its next waypoint where autonomous exploration resumed.

III. EXPERIMENTS

A. *Experiment I*

Experiment I reported in [8] and [9] compared a Control condition without alarms with two alarm conditions: Alarm in which all malfunctions were displayed on a status panel and FIFO which displayed alarms one at a time in the order in which they occurred. Because all failures were of the same difficulty the order in which they are serviced should make no difference so according to the attention scheduling hypothesis the FIFO and Alarm conditions should produce equivalent performance. The experiment followed a three condition repeated measures design comparing the conventional MrCS displays with MrCS augmented by alarm panels. Conditions were fully counterbalanced for Map/starting points and display with 5 participants run in each of the six cells

B. *Experiment II*

Experiment II reported in [11] and [9] extended the investigation begun in Experiment I by introducing multiple types of failures to allow a condition for which the schedule-aiding hypothesis would predict superior performance. Servicing the shortest job first (SJF) is a provably optimal policy for maximizing throughput in a queuing system [5]. An alarm system that displayed only the current failure with the shortest time to repair, therefore, should improve the performance of the human-multirobot system over the Alarm condition unless the unaided human is also following the same SJF policy.

Recoverable failures were categorized into 4 major types, based on the data for commonly occurring non terminal and field repairable failures for the Pioneer P3-AT [12]. Two of these, camera and map failures, involve loss of display due to communication difficulties. The third, teleoperation lag is a control problem found by [13] to significantly degrade operator performance. The fourth, "stuck", is a common condition in which a robot becomes entangled with obstacles. To resolve encountered failures, the operator needed to manually guide the robot from its current location to the next waypoint. Because each of the failure types imposed different difficulties for recovery, they took varying amounts of time to resolve.

C. *Results*

Data were analyzed using repeated measures ANOVAs comparing search and rescue performance between the control and two alarmed display conditions in Experiment I and between three alarmed display conditions in Experiment II. Where effects were observed pairwise comparisons were also conducted.

No difference was found on the overall performance measures of areas covered or victims found in either experiment, however significant effects were found, on measures relating to operator strategy and task performance in both experiments.

Table 1 summarizes the results from the experiments showing effects significant at the $p < .05$ level. Conditions shown within parentheses do not differ significantly from one another but are both significantly different from the third. Alarm and SJF conditions, for example, did not differ from one another but each had significantly fewer false positives than the FIFO condition.

In Experiment I we found that alerting operators to robots in need of interaction improved performance along a number of dimensions. The study compared a control condition without alerting with experimental conditions corresponding to the Alarm and FIFO conditions of Experiment II. While alerting was beneficial, FIFO which directed the operator to service a particular robot was less effective than the Alarm which allowed the operator to choose which alarming robot to service. This contradicts the predictions of the *attention scheduling hypothesis* which required that human attention be directed without loss of cognitive efficiency. The advantage for less constrained operators might be explained either by superiority of

Table 1. Summary of Data from Experiments I and II

TABLE 1: DATA SUMMARY	**Experiment I**	**Experiment II**	**Effects**
Primary Task (victim detection & marking) Performance Measures			
Area Covered	No Effect	No Effect	No Effect
N of Victims	No Effect	No Effect	No Effect
False Positives	Not Tested	FIFO > (Alarm, SJF)	FIFO > (Alarm, SJF)
Misses	Not Tested	Alarm > SJF > FIFO	Alarm > SJF > FIFO
Victim Delay	Alarm < (Control, FIFO)	Not Tested	Alarm < (Control, FIFO)
Select to Mark	Alarm < Control	Alarm < (SJF, FIFO)	Alarm < (SJF, FIFO, Control)
Secondary Task (failure detection & repair) Performance Measures			
Failures Resolved	Not Tested	(Alarm, SJF) > FIFO	(Alarm, SJF) > FIFO
Fault Detection	Alarm < Control		Alarm < Control
Full Task			
Neglect Time	FIFO > Control	No Effect	FIFO > Control
NASA-TLX	No Effect	No Effect	No Effect
Frustration	Not Tested	Alarm > (FIFO, SJF)	Alarm > (FIFO, SJF)

strategies of Alarm operators when allowed choice or operator difficulties in complying with automation that prescribed the robot to be serviced.

Experiment II partially supported the premise that operator attention can be directed to interaction with individual robots without degrading performance. Alarm performed slightly better than SJF on false positives, distance traveled, and failures resolved, but only for select-to-mark times did the difference approach significance. For the primary task of marking victims, FIFO participants proved slightly better, however, SJF participants were significantly superior to Alarm users yielding a balanced performance which was never poorest. The above results might have been due to the differences in allocation of attention. Within limited cognitive capacity of processing information, operators have to selectively dedicate attention to any of the "wanted" targets and filter out the irrelevant information simultaneously [17]. Alarm operators must devote time and attention to monitoring and selection of robots for servicing as well as the interaction leaving less available for the victim monitoring and marking tasks; whereas operators in the forced queue (Priority-/FIFO-Queue) conditions, by contrast, do not have to compete with monitoring and selecting robots to service, leaving more resources available for victim-related tasks, which may have led to the reversed results in unmarked victims among three conditions.

The FIFO-queue condition which directed operator attention suboptimally also led to the greatest loss of situation awareness as reflected in its longest Select-to-Mark victims times and lowest marking accuracy. This may have been exacerbated by the FIFO discipline which did not distinguish between distracting recoveries such as loss of track on map and brief interventions such as maneuvering around an obstacle. For the Priority-queue, the SJF discipline had not only the advantage of allowing operators to work primarily on briefer interventions thereby preserving SA, but by clustering similar types of failures increased opportunities for reducing the cost to switch between recovery strategies and sharing the similar cognitive procedures among failures. However, the Priority-queue operators may have simply devoted more of their time and attention to robot requests than operators using the less efficient FIFO because of their greater payoff, which could be observed from the higher rate of unmarked victims.

Performance on the primary victim detection and marking task was poorest in the Alarm condition with more misses and fewer false alarms suggesting operators may have been devoting less effort to this task. When they did see a victim, however, they were faster to select the robot and mark the victim then those using priority queues indicating better situation awareness. This advantage extended to the secondary task where Alarm users were faster to address and resolve faults. The performance improvements came at some cost, however, as indicated by the elevated frustration scale of the workload measure.

D. *Spatial Analysis*

Taken together these experiments fail to confirm the *attention scheduling hypothesis* as the FIFO and SJF interfaces that dictated the malfunctioning robot to be serviced led to decreased cognitive efficiency as reflected in poorer performance in direct comparisons. Two possible explanations were that:

1) *lack of volition in choice of robot to service led to inefficiencies due to task switching [14] or*

2) shifts of attention to potentially remote locations added extra costs in reacquiring situation awareness.

To address this question data were reanalyzed to examine spatial differences in shifts of attention between robots. Because operators were free to select robots to search for or mark victims as well as recover failures it was not possible to definitively determine the source of a shift in attention. We therefore examined all shifts of attention between robots on the assumption that shifts associated with the primary victim monitoring and marking tasks would be similar across conditions so that observed differences could be attributed to effects of the alarm conditions on the secondary error recovery task. The Euclidian distance between a selected robot and the previously selected one was used as a surrogate for differences in situation awareness based on the rationale that nearer robots were more likely to share landmarks and other common environmental features to support maintained situation awareness. The average distance of attentional shifts was defined as the sum of the distances between robots involved in selections divided by the number of selections.

Table 2 Spatial Distances from Experiment 1

Groups	Avg_Distance
Control (6 robots)	21.418
FIFO-queue (6 robots)	22.693
Alarm (6 robots)	23.078
ANOVA reveals an overall effect (F2,58=5.975, p=.004) T-test: Alarm > Control (p=.002) FIFO >Control (p=.020) No significant difference was observed between Alarm and FIFO conditions (p=.457)	

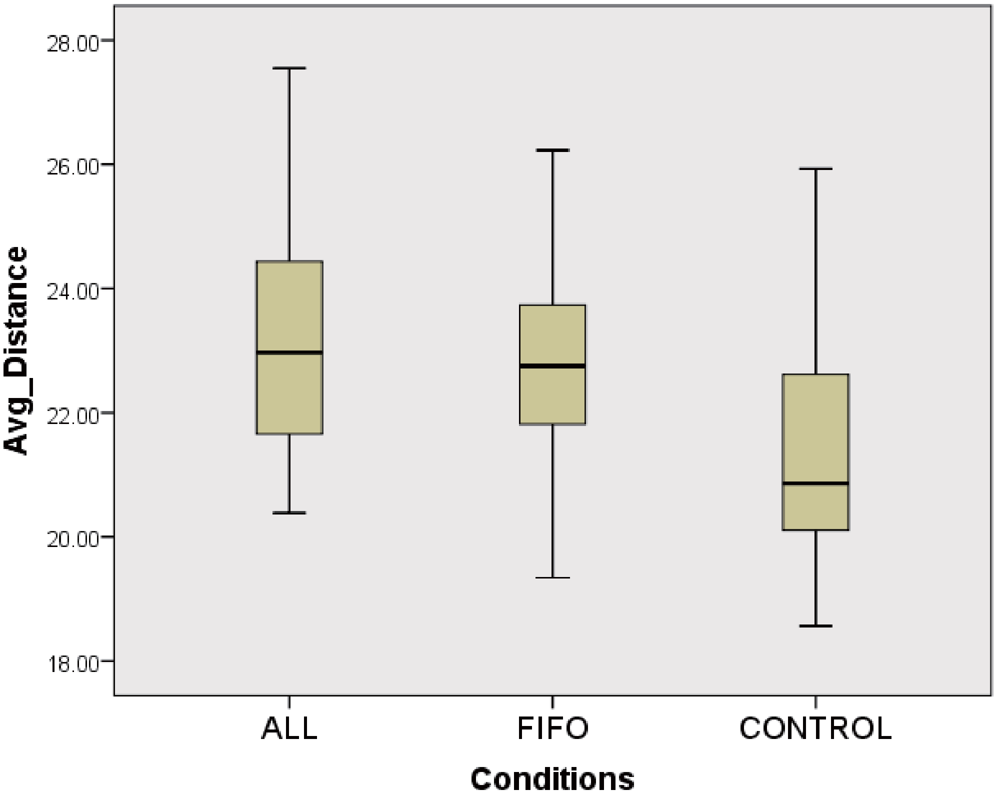

Fig. 3 Switching Distances between robots in Experiment I

Table 3 Spatial Distances between robots from Experiment II

Expt2_Group	**Avg_Distance**
SJF (8 robots)	43.034
FIFO (8 robots)	45.394
Alarm (8 robots)	42.613
No overall effect was observed in ANOVA (F2,58=1.783, p=.177). T-tests also revealed no significant difference among conditions	

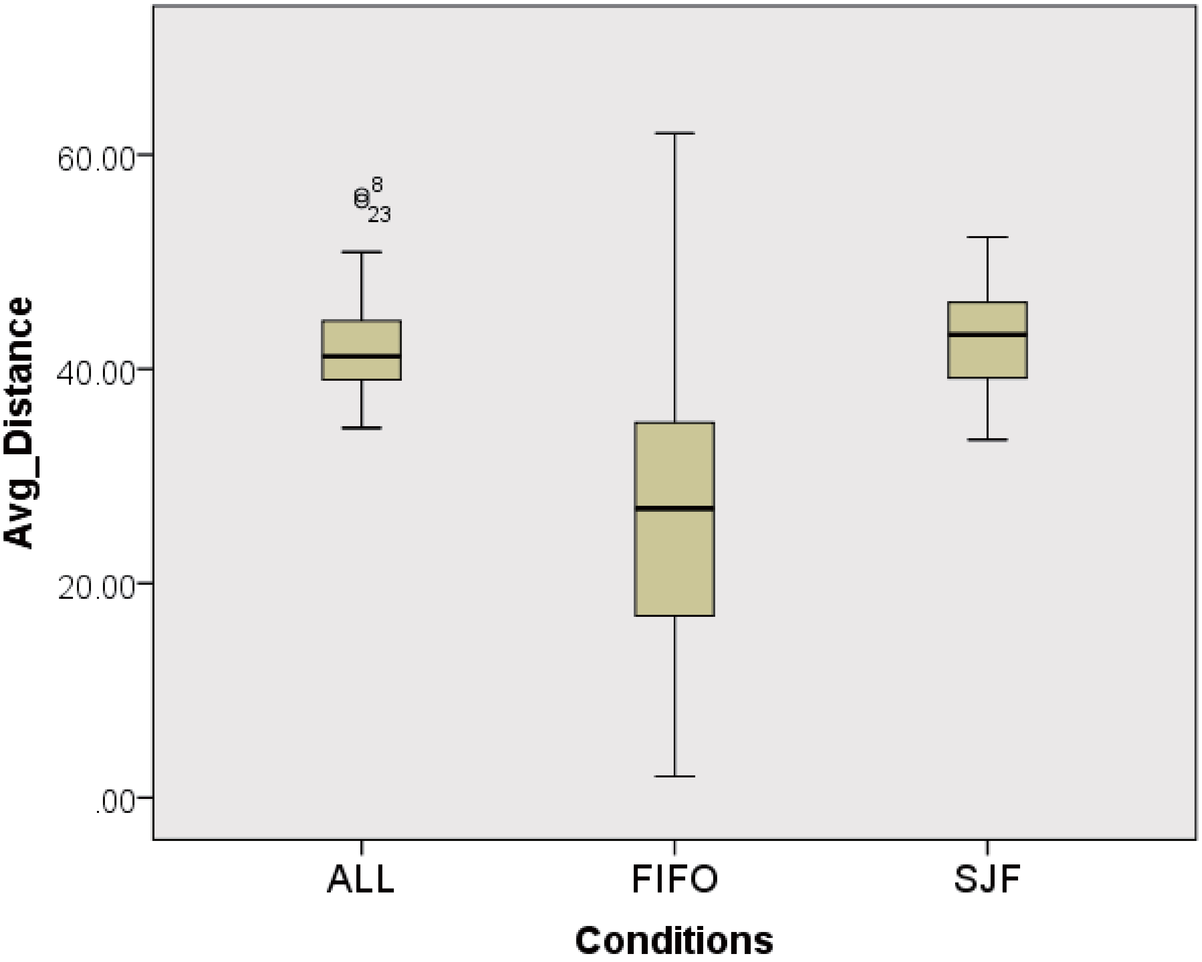

Fig 4. Switching Distances between robots in Experiment II

In Experiment I (Table 2, Figure 3) both alarm conditions (Alarm and FIFO) led to shifts between more distant robots than the control in which failures were detected through monitoring. The situation awareness explanation of the underperformance of forced queue conditions would have required that the FIFO condition in which the operator did not choose the robot to select would lead to longer distances. Instead, average distances were nearly identical indicating that the factor underlying the superiority of the Alarm condition was unrelated to the choice of nearby robots.

Experiment II (Table 3, Figure 4) was even less favorable to the situation awareness hypothesis finding no differences in distances among the three conditions. These results suggest that some other explanation of the advantages of access to all alarms is needed.

IV. DISCUSSION

Results of these two experiments do not fully support the *attention scheduling* hypothesis which is necessary if scheduling algorithms are to be used to improve human-multi-robot interaction. The *remote robot* explanation of the attention direction deficit that would have allowed incorporation of distances of attentional shifts into scheduling algorithms also does not appear tenable. There remain special conditions under which directed attention might provide advantages over freely available alarms.

In forced queue conditions operators receive an explicit recommendation for the robot to assist. Under extreme stress or time pressured tasks, humans tend to defer to automation and rely on the system for making decisions [16]. This increased compliance under high workload could be especially beneficial to system performance where optimal strategies such as SJF can be used to steer operator attention.

The study results are promising although not conclusive for the prospects of improving HRI performance through scheduling operator attention. The improvement of performance in queuing discipline leading to equivalent forced queue/Alarm performance in Experiment II shows that forced queue aiding can be effectively used by operators and might even lead to superior performance under more complex or stressful conditions where choice among robot requests becomes more difficult. The source of the deficit in directed attention when compared with freely available alarms, however, remains unsolved and a problem for future research.

REFERENCES

[1] Y. U. Cao, A. S. Fukunaga, and A. Kahng, "Cooperative mobile robotics: antecedents and directions," *Autonomous Robots*, vol. 4, no. 1, pp. 7-27, 1997.

[2] J. W. Crandall, M. A. Goodrich, D. R. Olsen, and C. W. Nielsen, "Validating Human–Robot Interaction Schemes in Multitasking Environments," *IEEE Transactions on Systems Man and Cybernetics Part A Systems and Humans*, vol. 35, no. 4, pp. 438-449, 2005.

[3] M. Scheutz and J. Kramer, "Reflection and Reasoning Mechanisms for Failure Detection and Recovery in a Distributed Robotic Architecture for Complex Robots," *Components*, no. 1, pp. 3699-3704, 2007.

[4] Y. Xu, T. Dai, K. Sycara, and M. Lewis, "Service Level Differentiation in Multi-robots Control," *System*, pp. 2224-2230,2010.

[5] M. R. Garey, D. S. Johnson, and R. Sethi, "The Complexity of Flowshop and Jobshop Scheduling," *Mathematics of Operations Research*, vol. 1, no. 2, pp. 117-129, 1976.

[6] S. Carpin, M. Lewis, J. Wang, S. Balakirsky, and C. Scrapper, "Bridging the gap between simulation and reality in urban search and rescue," *Robocup 2006 Robot Soccer World Cup X*, vol. 4434, pp. 1-12, 2007.

[7] S. Y. Chien, H. Wang, and M. Lewis, "Human vs . Algorithmic Path Planning for Search and Rescue by Robot Teams," *Human Factors*, vol. 54, no. 4, pp. 379-383, 2010.

[8] S.-Y. Chien, H. Wang, M. Lewis, S. Mehrotra, and K. Sycara, "Effects of Alarms on Control of Robot Teams," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2011, vol. 55, no. 1, pp. 434–438.

[9] Lewis, M., Chien, S., Mehortra, S., Chakraborty, N. & Sycara, K. (2014) Task switching and single vs. multiple alarms for supervisory control of multiple robots, International Conference on Human-Computer Interaction, (LNCS 8532), Springer International Publishing Switzerland, 499-510.

[10] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Human Mental Workload*, vol. 1, P. A. Hancock and N. Meshkati, Eds. North-Holland, 1988, pp. 139-183.

[11] S. Y. Chien, S. Mehrotra, M. Lewis, and K. Sycara, "Scheduling Operator Attention for Multi-Robot Control," Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2012), 2012, pp. 473-479.

[12] J. Carlson, R. R. Murphy, and A. Nelson, "Follow-up analysis of mobile robot failures," in *IEEE International Conference on Robotics and Automation 2004 Proceedings ICRA 04 2004*, 2004, vol. 5, no. April, pp. 4987-4994

[13] T. B. Sheridan, "Space teleoperation through time delay: review and prognosis," *IEEE Transactions on Robotics and Automation*, vol. 9, no. 5, pp. 592-606, 1993.

[14] A. Kiese., M. Steinhauser, M.Wendt M. Falkenstein, K. Jost, A. Philipp & I. Koch. (2010). Control and interference in task switching--a review, Psychological Bulletin, 36(5), 840-874.

[15] A. Kirlik, "Modeling strategic behavior in human-automation interaction: why an 'aid' can (and should) go unused.," Human Factors, vol. 35, no. 2, pp. 221-242, 1993.

[16] T. Inagaki, "Adaptive Automation : Sharing a Trading of control," in Handbook of Cognitive Task Design, E. Hollnagel, Ed. 2003.

[17] C. Billings, "Aviation Automation, Thesearch for the Human-Centered Approach," Lawrence Erlbaum, 1997.

[18] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation.," IEEE transactions on systems, man, and cybernetics. Part A, Systems and humans : a publication of the IEEE Systems, Man, and Cybernetics Society, vol. 30, no. 3, pp. 286-97, May 2000.