Teamwork in Controlling Multiple Robots

Fei Gao	M.L. Cummings	Luca F. Bertuccelli
Massachusetts Institute of Technology	Massachusetts Institute of Technology	Massachusetts Institute of Technology
77 Massachusetts Ave, Rm 33-309	77 Massachusetts Ave, Rm 33-309	77 Massachusetts Ave, Rm 33-309
Cambridge, MA, 02139	Cambridge, MA, 02139	Cambridge, MA, 02139
1-617-606-5522	1-617-252-1512	1-617-258-5046
feigao@mit.edu	missyc@mit.edu	lucab@mit.edu

ABSTRACT

Simultaneously controlling increasing numbers of robots requires multiple operators working together as a team. Helping operators allocate attention among different robots and determining how to construct the human-robot team to promote performance and reduce workload are critical questions that must be answered in these settings. To this end, we investigated the effect of team structure and search guidance on operators' performance, subjective workload, work processes and communication. To investigate team structure in an urban search and rescue setting, we compared a pooled condition, in which team members shared control of 24 robots, with a sector condition, in which each team member control half of all the robots. For search guidance, a notification was given when the operator spent too much time on one robot and either suggested or forced the operator to change to another robot. A total of 48 participants completed the experiment with two persons forming one team. The results demonstrate that automated search guidance neither increased nor decreased performance. However, suggested search guidance decreased average task completion time in Sector teams. Search guidance also influenced operators' teleoperation behaviors. For team structure, pooled teams experienced lower subjective workload than sector teams. Pooled teams communicated more than sector teams, but sector teams teleoperated more than pool teams.

Categories and Subject Descriptors

J.7 [**Computers in Other Systems**]

General Terms

Experimentation, Human Factors

Keywords

Multiple Robots, Teamwork, Communication, Urban Search and Rescue

1. INTRODUCTION

Enhanced autonomy makes it possible for one operator to control multiple robots. It releases an operator from manually controlling each robot and makes it possible to do tasks requiring monitoring, coordination, and complex decision-making. However, the required cognitive load for controlling multiple robots could

easily exceed that of a single operator, even with higher levels of automation. Teams are increasingly called upon to perform complex cognitive tasks that are less efficiently done by, and sometimes cannot be accomplished by, an individual. Although teamwork may impose extra workload related to coordination and communication, teams have the potential of offering greater adaptability, productivity, and creativity than any one individual can offer. Moreover, teams can provide more complex, innovative, and comprehensive solutions [5].

Unfortunately, the benefits of teamwork do not always occur naturally, and teams can fail for many reasons [14]. Factors such as poor combination of individual efforts, a breakdown in internal team processes (e.g., communication), and improper use of available information have been identified as potential sources of team failure [15]. In addition, when people collaborate with autonomous systems, system complexity inevitably increases, and automation can change the way people coordinate with each other [12]. To enable collaborative human-automation team interactions, we must therefore understand the nature of such teamwork, including outcomes, processes and dynamics.

Teaming difficulties in controlling multiple robots lead to the following research questions. First, is it possible to design automated decision support tools to improve performance and reduce workload when controlling multiple robots? Second, how does the organization and structure of team members affect performance, workload and communication?

The remainder of the paper proceeds as follows. In Section 2, we review related work. In Section 3, we introduce the research questions and describe an urban search and rescue experiment, including the testbed and experimental procedure. In Section 4, we describe the statistical results from the experiment. In Section 5, we summarize and discuss the results.

2. RELATED WORK

2.1 Allocating Attention across Multiple Robots

Controlling multiple autonomous robots is complex. When supervising a team of robots in a time-critical situation, the time and attention resources of operators are limited since operators are known to process complex tasks in a serial fashion [16]. As a result, there is a temporal opportunity cost associated with each task, and an implicit cost-benefit analysis that operators must perform in order to best allocate their attention across competing tasks. In the neglect tolerance model [3], an operator interacts with one robot for a period called Interaction Time (IT), then neglects it for a period called Neglect Time (NT) to interact with other robots before the first robot must be revisited to maintain performance. The number of independent homogenous robots that

a single operator can control is calculated by NT/IT+1. Further studies showed that the number of robots an operator can control is affected by other factors such as the nonlinear increasing complexity [9] and switching cost between robots [2].

Previous work in automated visual search task allocation for single operator-multiple unmanned vehicle environments by Bertuccelli et al. [1] has shown that automated search guidance can improve operator performance in terms of overall mission probability of detection and lower workload by influencing switching times. This form of search guidance was hypothesized to be beneficial also in the context of team scenarios, where resources were distributed across operators, who could benefit from recommendations for when to switch to new search tasks.

2.2 The Role of Team Structure
Team structure is another important factor hypothesized to affect team effectiveness for the search and rescue setting [8]. Team structure can be described as the work assignment and communication architecture. Work assignment is the "manner in which the task components are distributed among team members" [13]. How the team is structured is closely related to communication, coordination and team performance.

The team structure that is suitable for a specific scenario largely depends on the task characteristics and resources available [11]. For a team of operators working together with multiple homogeneous unmanned vehicles, two possible ways to organize the vehicles are as Sectors or as a Shared Pool [7]. In the Sector condition, each operator controls a portion of all the vehicles. In the Shared Pool condition, operators share the control of all the robots and service them as needed. Sector assignment, which is how modern day air traffic control is architected, can reduce the number of robots the operator must monitor and control. However, the Shared Pool condition offers a more flexible scheduling advantage of load balancing since any operator in the team can service any robot as needed. In addition, for monitoring applications, the Shared Pool offers a redundant observer advantage, such that a second operator with partially overlapping perceptual judgments may detect victims missed by the first operator.

Coordinated action lies at the heart of effective team performance. Communication is an important method of explicit coordination. It also relates to building an accurate understanding of team members' needs, responsibilities, and expected actions [11]. Communication also requires cognitive and attention resources, and may be hindered because the environment has become stressful and team members focus on their individual tasks rather than on how those tasks affect other team members [15].

Previous research by Lewis and Wang et al. [8] investigated the effect of autonomous path planning versus manual control and team structure in a Urban Search and Rescue (USAR) setting. They found automating path planning improved system performance but it may weaken situation awareness. For team structure, no significant difference on performance was found, but teams that shared the control of all robots had slightly lower workload.

Using a similar experiment setting, we investigated similar issues including 1) Understanding the role of automated search guidance on operator and team performance, workload and communication when controlling multiple robots, and 2) Understanding the role of team structure on overall mission performance, subjective workload, and working process. This experiment is detailed in the next section.

3. EXPERIMENT DESIGN

3.1 Testbed
USARSim, a robotic simulation performing Urban Search and Rescue (USAR) tasks [10], was used to provide the underlying simulation for the testbed. MrCS (Multi-robot Control System), a multi-robot communications and control infrastructure with an accompanying user interface was used as the control interface. MrCS provides facilities for starting and controlling robots in the simulation, displaying camera and laser range finder output, and supporting inter-robot communication through Machinetta, a distributed multi-agent system developed at Carnegie Mellon University [4]. Figure 1 shows the elements of the MrCS displayed on a dual display computer. Thumbnails of robot camera feeds are shown on the left screen. A video feed of interest is on the top left of the right screen. Under the video feed, a GUI element in the bottom left allows teleoperation and camera pan and tilt. The right shows the current area map and allows operators to mark the location of victims.
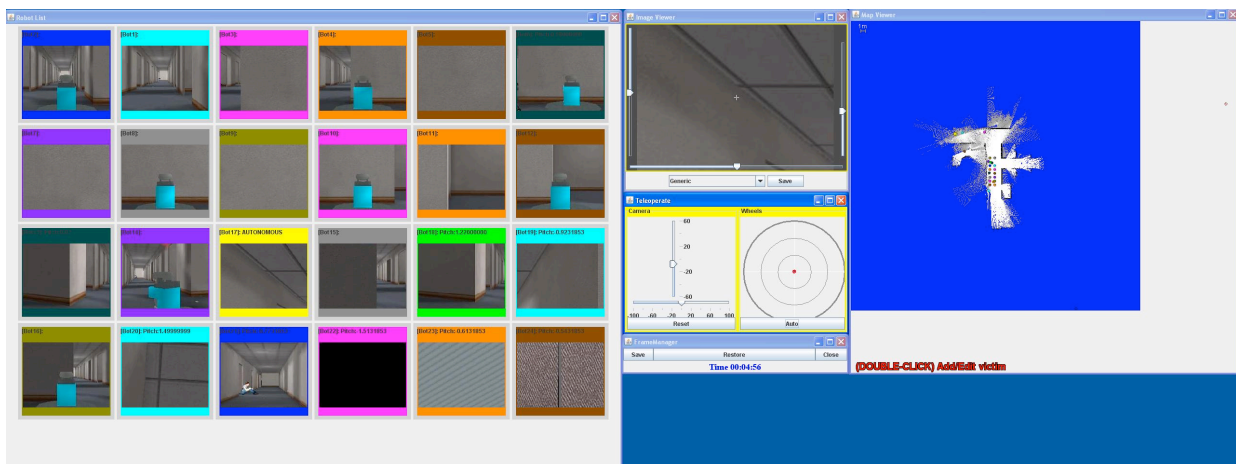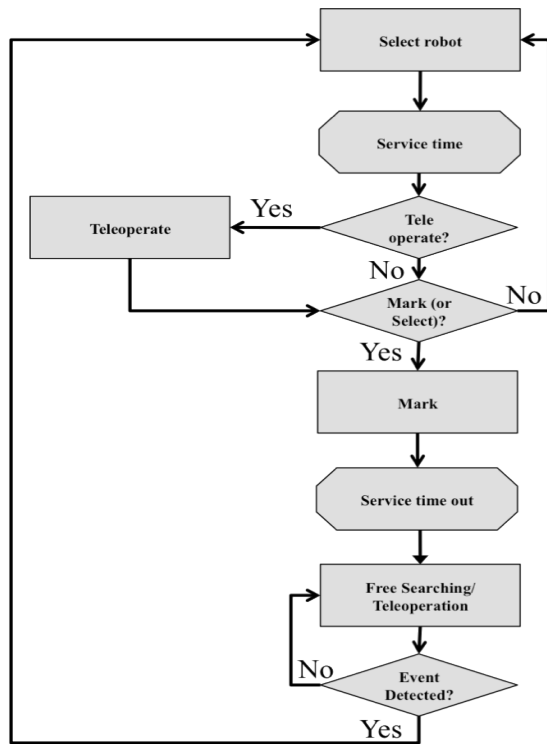

Figure 1: Interface for operating vehicles

Figure 2: Operation procedure

In MrCS, each robot is capable of updating a map, planning their routing and sending back video feed to operators. The operators' tasks were to explore the environment and identify as many positions of victims as possible. There was little interdependency between robots. In this experiment, robots were started in different regions and explored the environment automatically. The operators guided the robots in the environment in order to find the victims. The general workflow of a single operator is shown in Figure 2. When a victim appeared in the camera of a robot and was detected by the operator, the operator's task was to select the robot, teleoperate the robot to bring the victim back into the camera view and mark the location of the victim on the map. This was the time the operator devoted to serve the robot, labeled as service time/service time out in Figure 2. After that, the operator continued monitoring all the robots and guided the robots to explore the environment. Most of the time robots were moving around using autonomous path planning, and the operator only needed to monitor the thumbnails of video feeds. It also happened that the operator used teleoperation to manually control the robots to send them to a specific unexplored place. This free searching/teleoperation period stopped until a new victim appeared in a camera view, and the operator selected this robot to start a new task.

3.2 Independent Variables

The goal of this experiment was to investigate the effect of search guidance and team structure on operator and team performance, workload and communication in an Urban Search and Rescue task with multiple robots.

A new automated search assistant was designed and inserted into the testbed that recommended an appropriate time for operators to interrupt their current searches and investigate other imagery. The three types of automated schedule assistants were as follows:

1) In the Off (O) condition, participants received no decision support and each participant was tasked to search and mark the victims, and teleoperate a selected robot when needed.

2) The Suggested (S) condition gave a notification when the operator spent more than 30 seconds on a robot, and recommended that the operator move on to another robot.

3) The Enforced (E) condition gave a notification when the operator spent more than 30 seconds on a robot and switched to another robot automatically after 5 seconds.

Thirty seconds was chosen as the threshold criteria based on previous study [1, 8, 9] and pilot test of the experiment. In previous studies on visual search tasks [1], the possibility of finding a target was shown to decrease as more time was spent on the visual search task. The probability was estimated to be 0.8 for 26 seconds spent on searching. In another experiment for USAR tasks [8], the mean time from a robot being selected to a victim being marked under autonomous control was approximately 35 seconds. We selected 30 seconds as the threshold so that an operator was given reasonable amount of time to finish the task if a victim was successfully located and was prevented from spending too much time on a low probability search task if the operator failed to locate the victim. This threshold was validated in pilot tests as well.

In this experiment, operators were grouped into two types of team structures:

1) In Sector (S) teams, each participant controlled half of all the robots, for a total of 12 robots. Locations of their teammates' robots were shown on the map, but video feed from their teammates' robots could not be seen.

2) In Pool (P) teams, two operators shared the control of all the robots. They were able to see the video feed of all robots and control any robot not under control by a teammate.

3.3 Dependent Variables

Dependent variables included task performance metrics, subjective workload, operator measures and communication as team measure. Task performance metrics included number of victims found, number of deletes, number of errors in marking victims, number of victims missed, and percentage of area explored. Number of victims found was evaluated based on the distance between the mark position and real position of the victim. A mark was correct if the distance was less than one meter, otherwise it was an error. Number of deletes was the number of marks deleted by the operator. Deletes happened when a victim was not accurately marked previously or was marked for more than once, and this measure was related to the accuracy of marking victims. Subjective workload ratings were obtained through the NASA-TLX [6], which measures six sub dimensions. Operator measures included display-to-mark time, select-to-mark time, teleoperation frequency, duration and total teleoperation time. Communication was evaluated based on total communication time. All the dependent variables were summarized in Table 1 along with their definitions.

Table 1: Dependent Variables

Category	Dependent Variables
Task performance metrics	Found: number of victims mark in the correct position
	Error: number of marks in the wrong position
	Deletes: number of marks deleted
	Missed: number of victims that appeared in the camera but were not marked
	Percentage of area explored: area explored/total area
Workload	NASA-TLX rating
Operator Measures	Teleoperation duration: length of teleoperation period before marking victim or robot selection
	Teleoperation frequency: number of teleoperation
	Total teleoperation time: total amount of time spent on teleoperation
	Display-to-mark time: time from victim appearing in the camera to being marked
	Select-to-mark time: time from robot selection to victim being marked
Team Measure: Communication	Communication time: total time spent communicating with team member

3.4 Procedure

The experiment began with a 15-minute training session prior to three 25-minute test sessions. Participants were tested in groups of two to facilitate teamwork in performing the USAR task. Each participant controlled 12 robots individually in the Sector condition, or all 24 robots in the Shared Pool condition. Each pair of participants performed all three search guidance conditions. The three conditions were randomized and counterbalanced to limit any learning effect. Audio and screen recordings were collected during the experiment. Subjective workload was rated using the NASA-TLX at the end of each session.

3.5 Participants

A total of 48 participants, aged 19 to 47 years old, participated in the experiment. The average age was 26.6 years, with a standard deviation (SD) of 5.5. Among them, 19 were female and 29 were male. Thirty-three of the participants were undergraduate or graduate students, and 15 had other occupations. Twenty-two of the participants did not play video games regularly. The average time playing video game per week for the remaining 26 participants was 4.1 hours (SD = 4.9). The correlation between hours spent on video games and average individual performance was not significant ($p = 0.354$).

4. RESULTS

Data logged during the experiment were post processed to obtain performance and process data. The criterion for a successfully marked victim was that the position of the mark was within one meter of the true position of the victim, which was the same criterion as in the study of Lewis et al. [8] In order to find the time when a victim appeared in the camera, we drew the visible areas

of all victims using ray tracing. If the robot was in the visible area for a victim, and its field of view contained the victim, this victim was declared visible on this robot's camera. By calculating these quantities, we were able to obtain the number of victims missed, and record the display-to-mark time. All the other dependent variables were calculated directly from the user interaction log.

The results were analyzed based on the three experiment sessions from four aspects: task performance metrics, subjective workload, operator measures and communication as team measure. Data in the training session was not included in the analysis.

4.1 Task Performance Metrics

The dependent variables of number of victims found, number of errors and percentage of area explored were analyzed using analysis of variance (ANOVA). The number of marks deleted and number of victims missed were analyzed using nonparametric tests since they did not satisfy the ANOVA assumptions of normality and/or homogeneity. An alpha of 0.05 was used in the analysis, and a p-value between 0.05 and 0.1 was defined as marginal significant.

For number of victims found, no significant effect was found for either of the primary independent factors. Session order, a secondary independent variable that represents the order the three search guidance conditions were performed, was shown to have a significant effect on number of victims found (Figure 3). After removing data from the first session of each experiment due to a learning effect, the interaction between team structure and search guidance showed a marginally significant effect ($F=2.591$, $p=0.087$). The Sector team had better performance without automated search guidance (Mean = 22.8, SD =3.73), while Pool teams worked better with enforced search guidance (Mean = 22.6, SD = 4.43), as shown in Figure 4.
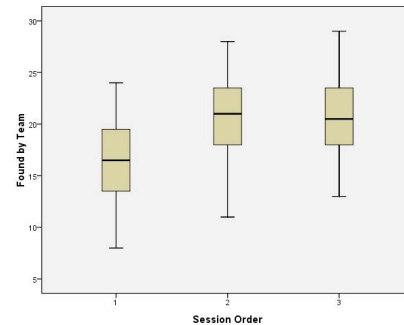

Figure 3: Number of victims found versus session order
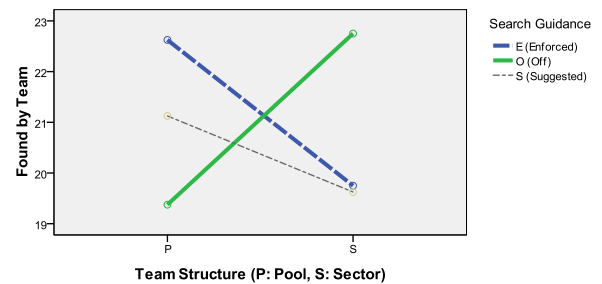

Figure 4: Interaction effect of team structure and search guidance on number of victims found

Although number of victims found was evaluated at the team level, operators in teams have difference on victims found by individual team members. Operators in Pools teams have significant larger difference on individual performance than in Sector teams (p=0.036), as shown in Figure 5.
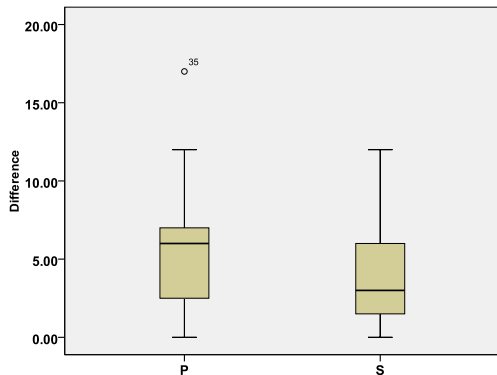

Figure 5: Difference on number of victims found by individual team members

For number of marks deleted (Figure 6), a marginally significant effect was found for team structure (p=0.066). Search guidance did not demonstrate a significant effect. Pool teams (Mean = 8.3, SD = 5.49) deleted more than Sector teams (Mean = 6.0, SD = 3.02), indicating Pool teams corrected themselves more often. Deletes also happened when a victim was marked by both operators, which could only happen in Pool teams. One interesting note is that the standard deviation of deletes with no search guidance of Sector teams (SD=1.946) is much lower than in the other conditions, as shown in Figure 6. This means the deleting behavior was more consistent across different operators under the Sector condition, and operators were less affected by their teammate.
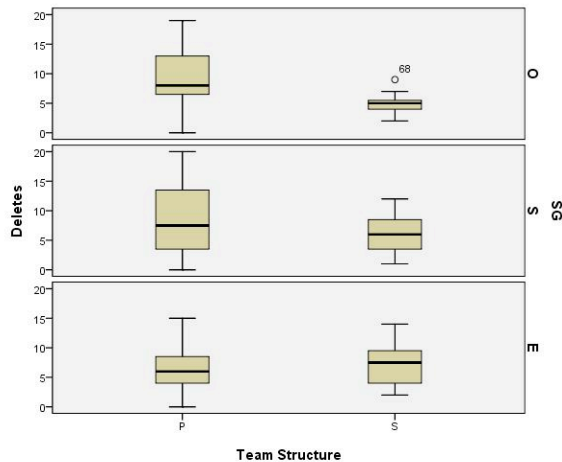

Figure 6: Number of deletes

No significant main effects were found for number of errors, number of victims missed and percentage of area explored across either the automated search guidance or structure independent factors. Moreover, these performance indicators were correlated with each other. Area explored was positively correlated with number of victims found by team (r=0.470, p<0.001). This is because when larger area was explored, there was a larger chance of finding a new victim. The number of victims found was expectedly negatively correlated with number of victims missed (r=-0.514, p<0.001) and number of errors (r=-0.691, p<0.001). In other words, teams that found more victims made fewer errors and missed fewer victims.

4.2 The Effect on Subjective Workload

Subjective workload using NASA-TLX was analyzed using non-parametric tests. Box plots of subjective workload under different conditions are shown in Figure 7. Mann-Whitney tests for the effect of team structure showed a significant effect on workload (p=0.042). Operators in Pool teams demonstrated lower workload on average than those in Sector teams. When analyzing each dimension of workload (mental demand, physical demand, temporal demand, performance, effort and frustration) separately, team structure had a significant effect on effort (p=0.032) and frustration (p=0.005). This was consist with previous study [8], in which a slight advantage in workload was observed favoring the Pool structure. One reason may be that in the sector team conditions, operators felt a lot of pressure because if they missed a victim, no one backed them up. Furthermore, in Pool teams, it is possible to balance the workload according to operators' individual abilities. When one operator was better at finding victims, it is possible he/she could share the burden of the less skilled teammate and did not report excessive workload. In Sector teams, all the work was split equally regardless of ability. We then analyzed subjective workload at the team level by taking the maximum, minimum and average of individual team members' subjective workload. Results show that maximum workload of the team members in Pool teams is significantly lower than in Sector teams (p=0.012), while average workload and minimum workload did not significantly differ. This result, combined with the significantly larger difference on individual performance in terms of number of victims found within Pool teams (Figure 5), suggests a workload balancing process or back-up behavior in Pool teams.
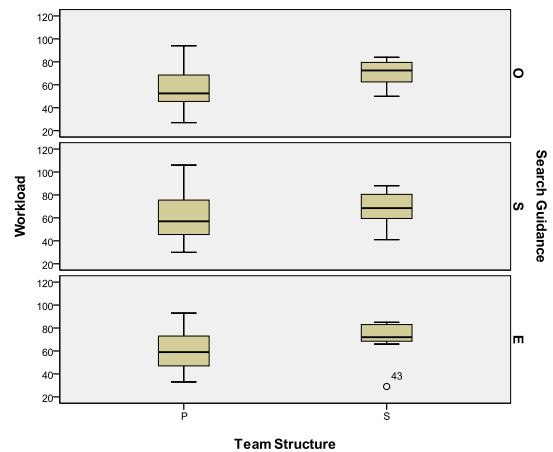

Figure 7: Subjective workload

4.3 The Effect on Operator Measures

The operator measures included the mean duration of teleoperation, frequency and total time of teleoperation, display-to-mark time and select-to-mark time. Total time of teleoperation was analyzed using ANOVA, while all the others were analyzed using nonparametric analysis because normality assumptions were not satisfied.
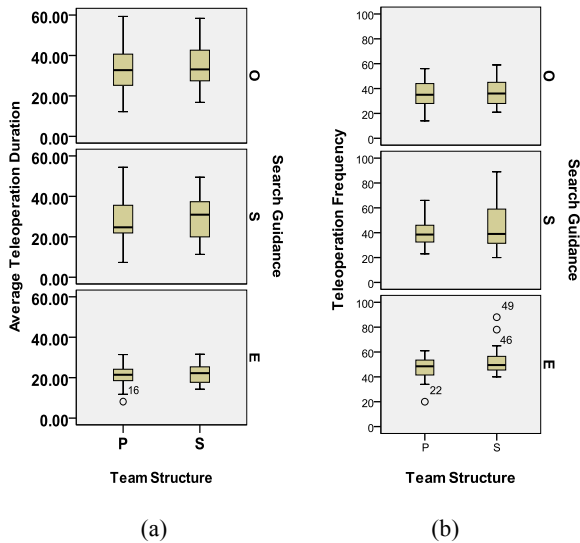
(a)   (b)

Figure 8: (a) Average teleoperation duration (seconds) and (b) teleoperation frequency

Search Guidance was found to have a significant effect on the duration (p<0.001) and frequency of teleoperation (p<0.001). As shown in Figure 8, under the enforced condition, the duration was shorter (Mean=21.7, SD=5.31, p=0.001) and the frequency was higher (Mean=50, SD=10.77, p<0.001) than under the suggested condition. Under the suggested condition, the duration was shorter (Mean=29.3, SD=11.55, p=0.025) and the frequency was higher (Mean=42.19, SD=13.79, p=0.045) than without search guidance. No significant effect of team structure was found for duration (p=0.687) and frequency of teleoperation (p=0.184).

For total time of teleoperation (Figure 9), team structure had a significant effect (p=0.001). Sector teams (Mean = 1166.7, SD = 194.20) spent more time on teleoperation than Pool teams (Mean = 1055.4, SD = 236.66) on average. Search guidance had a marginal significant effect (p=0.078). Operators expectedly spent more time on teleoperation when there was no search guidance condition, followed by the suggested condition and the enforced condition.
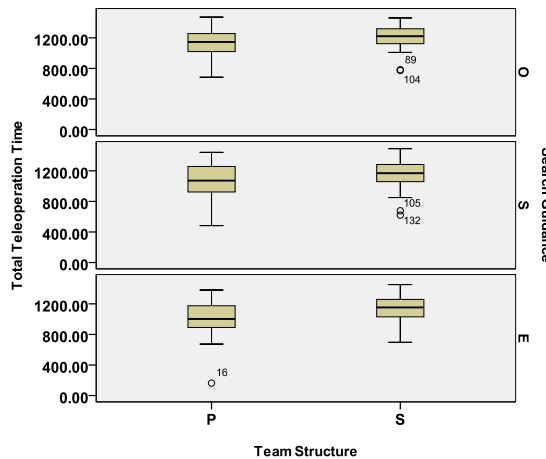

Figure 9: Total teleoperation time (seconds)

No significant main effect was found for display-to-mark time from either team structure (p=0.368) or search guidance (p=0.309). However, there was an interaction effect of search guidance and

team structure. As shown in Figure 10, search guidance significantly affected mean display-to-mark time in Sector teams (p=0.024). The teams in suggested search guidance condition had the lowest mean display-to-mark time (mean=88. 0s, SD=58.9s), followed by no guidance (mean=103.2s, SD=59.1s) and enforced guidance (mean=128.6s, SD=70.8s). This indicates that suggested search guidance helped the operator notice and mark victims faster when they appeared in the camera, which is very important for such a time-critical task environment. The increase in time in enforced guidance condition may be due to the interruption in the current operation and extra time to regain situation awareness. In Pool teams, the effect of search guidance was insignificant, suggesting that display-to-mark time is affected by a more efficient team process, i.e., one team member can start working on a robot with a victim in view when the other is busy.
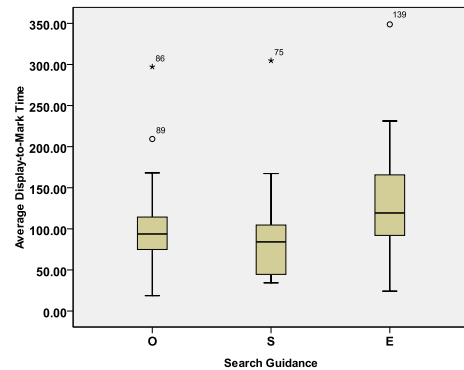

Figure 10: Average display-to-mark time (seconds)

For select-to-mark time (Figure 11), search guidance was found to have a significant effect (p=0.000), but no significant effect was found for team structure (p=0.331). For search guidance, it shortened the time to finish a task as expected. Operators may speed up their operations when a notification popped up. The change was more meaningful for the difference between Off condition and Suggested condition, because under enforced condition, this change may result from reselection of the same robot when the system switched to another robot automatically. For team structure, it did not have a significant effect on the time operators spent to serve a single robot. In other words, it didn't affect the task level operations.
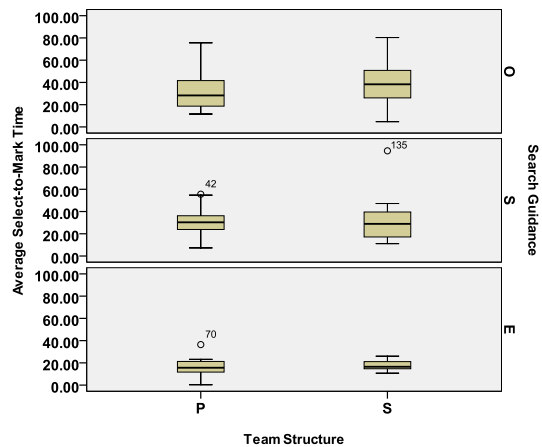

Figure 11: Average select-to-mark time (seconds)

Since teleoperation played an important role in operators' behavior, we analyzed the correlation between total teleoperation

time of each operator, subjective workload, the number of victims found by each individual operator and team performance. Total teleoperation time was not significantly correlated with the overall subjective NASA TLX workload (r=-0.004, p=0.965). However, it was significantly moderately correlated with the dimension of mental demand (r=0.372, p=0.001), so not surprisingly, using teleoperation was difficult.

There was a significant but weak correlation between total teleoperation time and number of victims marked by each operator (r=0.189, p=0.023). The total teleoperation time for the team of two operators was combined and analyzed with relation to team performance. It was found that total teleoperation time more strongly positively correlated with percentage of area explored (r= 0.431, p=0.000). Since autonomous path planning was used, the percentage of area explored largely depended on how much the operators changed the behavior of robots by manually redirecting them. In many situations, operators were not satisfied with the performance of autonomous path planning, so they controlled the robots using teleoperation to send them to the places they wanted the robots to go.

Lastly, total teleoperation time moderately correlated negatively with number of deletes (r=-0.331, p=0.005). With more teleoperation, operators could position the robots nearer to victims, which resulted in an increase in accuracy, thus a decrease in correcting behavior. No correlation was found between teleoperation time and number of victims found, number of errors or number of victims missed.

4.4 The Effect on Communication

During the experiment, participants were allowed to talk with each other. In such a high workload scenario, almost all the communication was mission-related. Teams who communicated discussed what strategies to use when exploring the area, updated his or her status with the teammate, requested his or her teammate's status or shared experiences with controlling the robots. In contrast, some teams did not communicate at all. A nonparametric analysis of the time spent on communication shows that team structure had a significant effect on communication time (p=0.004). As shown in Figure 12, Pool teams (Mean = 177.7, SD = 198.74) expectedly communicated more than Sector teams (Mean = 53.44, SD = 80.97), on average. Search guidance did not have a significant effect on communication time (p=0.865). Session order was found to have a marginally significant effect (p=0.098).
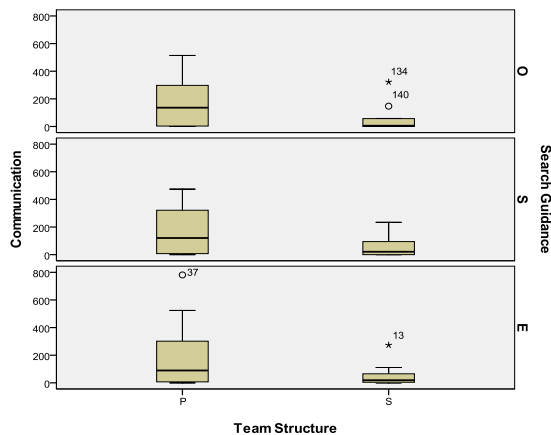

Figure 12: Communication time (seconds)

Further analyses on the correlation between communication with team performance and subjective workload revealed that communication time was moderately negatively correlated with errors (r=-0.309, p=0.008). In other words, teams that communicated more tended to make fewer errors. This negative correlation between communication time and number of errors exists for Pool teams but not for Sector teams. This result, combined with the larger number of deletes in Pool teams, suggests that mutual performance monitoring exists such that team members correct themselves and their teammates, facilitated by communication. Communication time was marginally significantly correlated with number of victims marked by teams (r=0.202, p=0.088). No significant correlation was found between communication time and number of deletes, number of victims missed, or subjective workload ratings.

5. DISCUSSION AND CONCLUSION

Although automated search guidance did not improve performance, it did not decrease it either. For teams under the Pool condition, it appeared beneficial to have an enforced search guidance to keep them from fixating on a single robot. In Sector teams, suggested search guidance helped operators mark victims faster when they appeared in the cameras. Furthermore, adding automated search guidance to the system changed operators' behavior in terms of teleoperation and select-to-mark time. By putting a threshold on time, search guidance shortened the episode of operation on teleoperation and select-to-mark. This changed behavior may be because operators hastened their operations when there was a notification, and partly due to reselection under Enforced condition. So in this experiment, spending less time on tasks did not significantly affect task performance.

When answering open-ended questions at the conclusion of the experiment, some participates said it was good to have a notification but the enforced change was annoying. In addition, although all the robots were the same, their importance was different according to their position in the environment. Robots near an unexplored area were more valuable than those in a thoroughly searched area. In this case, improving the automated search guidance by including location-related information and encouraging the operator to investigate unexplored area could increase the overall benefit of such a decision support tool.

Pool team structure was shown to be better than Sector structure in terms of experiencing lower workload but similar task performance. Although Pool teams corrected themselves more often, they had lower subjective workload. Pool teams also communicated more while Sector teams teleoperated more. Pool teams appeared to better balance workload among team members according to abilities. This conclusion was supported by the lower maximum subjective workload and larger difference on individual performance in Pool teams. This suggests the reduced subjective workload in the Pool condition was because teammates could provide back up if needed.

Teams with different structures did not show a significant difference in their task performance, except for number of deletes. The reason may be that they used different strategies to cope with the situation. Sector teams communicated less and teleoperated more, with the opposite generally true for Pool teams. One explanation is that in the Sector condition, operators had more time to spend on teleoperation since communication was not explicitly needed. The increased time spending on teleoperation extended the area they explored and increased the chances of finding new victims. On the other hand, the shared control of

robots promoted communication in teams under the Pool structure, which was also good for task performance since teams with more communication tend to make fewer errors. The reason may be that they corrected each other via communication, which led to fewer errors.

Overall in this study, operators tended to be robust to whatever structure they were assigned. Moreover, improvements could be made to the automated search guidance tool as the threshold for suggesting or enforcing a change was rudimentary. With a more context-sensitive tool, possibly based on an individual operator's strengths or limitations, it is possible that greater performance improvements could have been seen. In addition, an experimental confound that necessarily exists is the fixed number of victims in the environment, which bounds the performance on number of victims found. This also means the chance of finding a new victim decreases over time. The advantage of a team that is faster in marking victims over slower teams may not be shown on ultimate team performance. From this point, process measures are important and should be further investigated. We leave these issues as areas of future research.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Bertuccelli, L. F., Pellegrino, N. et al. 2010. Choice modeling of relook tasks for UAV search missions. *American Control Conference (ACC)*, 2010.

[2] Crandall, J.W. and Cummings, M.L. 2007. Attention Allocation Efficiency in Human-UV Teams AIAA Infotech@Aerospace Conference, Rohnert Park, CA, 2007.

[3] Crandall, J.W., Goodrich, M.A., Olsen, D.R., Jr. and Nielsen, C.W. 2005. Validating human-robot interaction schemes in multitasking environments. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, *35* (4). 438-449.

[4] Farinelli, A., Scerri, P. and Tambe, M. 2003. Building large-scale robot systems: Distributed role assignment in dynamic, uncertain domains. iAAMAS'03 Workshop on Resources, role and task allocation in multiagent systems.

[5] Gladstein, D. L. 1984. Groups in Context: A Model of Task Group Effectiveness. *Administrative Science Quarterly* 29(4): 499-517.

[6] Lewis, M., Polvichai, J. et al. 2006. Scaling-up Human Control for Large UAV Teams. *The Human Factors of Remotely Piloted Vehicles*. N. Cooke, Ed. New York, Elsevier: 237-250.

[7] Hart, S.G. and Staveland, L.E. 1998. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. in Hancock, P.A. and Meshkati, N. eds. *Human Mental Workload*, North Holland Press, Amsterdam.

[8] Lewis, M., Wang, H. et al. 2011. Process and Performance in Human-Robot Teams. *Journal of Cognitive Engineering and Decision Making* 5(2): 186-208.

[9] Lewis, M., Wang, H., Chien, S.Y., Velagapudi, P., Scerri, P. and Sycara, K. 2010. Choosing Autonomy Modes for Multirobot Search. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *52* (2). 225-233.

[10] Lewis, M., Wang, J. and Hughes, S. 2007. USARSim: Simulation for the Study of Human-Robot Interaction. *Journal of Cognitive Engineering and Decision Making*, *1* (1). 98-120.

[11] Macmillan, J., Entin, E. E. et al. 2004. Communication Overhead: The Hidden Cost of Team Cognition. *Team Cognition: Understanding the Factors That Drive Process and Performance*. E. Salas and S. M. Fiore, Ed., American Psychological Association (APA).

[12] Mouloua, M. and Parasuraman, R.1996. *Automation and Human Performance: Theory and Applications*, CRC Press

[13] Naylor, J. C. and Dickinson, T. L. 1969. Task structure, work structure, and team performance. *Journal of Applied Psychology* 53(10).

[14] Salas, E. and Fiore, S. M. 2004. *Team Cognition: Understanding the Factors That Drive Process and Performance*, American Psychological Association (APA)

[15] Salas, E., Sims D. E., et al. 2005. Is there a Big Five in Teamwork? *Small Group Research* 36(5): 555-599.

[16] Wickens, C. D. and Hollands, J. G. 1999. *Engineering Psychology and Human Performance*, Pearson Education.