2010-10-01

Beyond Robot Fan-Out: Towards Multi-Operator Supervisory Control

Michael A. Goodrich
mike@cs.byu.edu

Yisong Guo

Jonathan M. Whetten

Beyond Robot Fan-Out: Towards Multi-Operator Supervisory Control

Jonathan M. Whetten, Michael A. Goodrich, and Yisong Guo

Computer Science Department

Brigham Young University

Provo, Utah, USA

mike@cs.byu.edu

Abstract—**This paper explores multi-operator supervisory control (MOSC) of multiple independent robots using two complementary approaches: a human factors experiment and an agent-based simulation. The experiment identifies two task and environment limitations on MOSC: task saturation and task diffusion. It also identifies the correlation between task specialization and performance, and the possible existence of untapped spare capacity that emerges when multiple operators coordinate. The presence of untapped spare capacity is explored using agent-based simulation, resulting in evidence which suggests that operators may be more effective when they operate at less than maximum capacity.**

Index Terms—**Human-robot interaction, multi-operator supervisory control, fan-out**

I. Introduction

Many problems can benefit from having multiple robots including, lunar construction, search and rescue, and hazardous materials handling. A common theme in these applications is the need to minimize the number of humans required to manage the robots which, in turn, implies the desire to maximize human effectiveness. In the limit, this desire is often codified as maximizing the number of robots that a single human can manage, that is, maximizing fan-out [1].

Some question the utility of myopically maximizing fan-out, pointing out that the real desire is to maximize human effectiveness while maximizing team performance [2]. One way to shift focus from maximizing fan-out to maximizing team performance is to explore multi-operator supervisory control (MOSC) where multiple operators share responsibility to supervise multiple remote robots. In this paper, we contribute to the MOSC literature by identifying two important constraints on MOSC performance, task saturation and task diffusion, and by identifying how spare capacity and specialization can contribute to MOSC performance. Task saturation, task diffusion, spare capacity, and specialization support the assertion that myopically seeking to maximize fan-out may not maximize team performance.

The paper approaches MOSC research using two complementary research approaches: a human factors experiment and an agent-based simulation. We present results from a human factors experiment that identifies how task saturation and task diffusion naturally constrain MOSC performance. Results from this experiment also indicate the importance of specialization of roles among operators, and suggest that having multiple operators may produce spare capacity in these operators. We then use agent-based simulation to further explore how spare capacity affects MOSC performance, concluding that spare capacity is desirable and can be used to outperform teams that operate at or near maximum capacity.

II. Related Literature

The number of robots that can be controlled by a single human has been modeled using fan-out [3]. Cummings et al. extended Olsen's original model [1] to include wait times and performance metrics to try to model the "optimal" level of fan-out for a given set of task constraints [4]. Goodrich et al. extended Olsens model to include switch costs and heterogeneous robots [5]. In practice, fan-out ranges from one robot being controlled by two to three operators [6] to 12-16 robots being controlled by a single operator [7]. Cummings et al. [7] indicate that the level of autonomy of control, navigation, and payload management is key to higher levels of fan-out. Dudek at al. state that coordination costs imply that homogeneous robot teams grow slower than $1 \times N$ in the number of robots N, but autonomy-supported coordination may be able to overcome these efficiency limits.

Complementing autonomous robot coordination, it is possible to add multiple operators, an approach that can benefit from research in human teams. In his seminal work, Steiner [8] explores team process and limiting factors on productivity. Steiner provided a formula to calculate the productivity of groups, assuming that tasks can be divided among group members:

Actual Productivity = Potential Productivity - Losses Due to Faulty Process.

Losses in efficiency from groups include *production blocking* [9], *social loafing* or *social impairment* [10], *cognitive interference* [11], and *communication effort* [12]. Although the losses due to inefficiencies in groups can be large, this paper emphasizes potential productivity.

At least two specific group process gains have been noted in the literature. The *social facilitation effect* is based on the observation that performing a task in the presence of others can cause performance to significantly increase [13]. The *assembly bonus effect* occurs when group interaction combines members' knowledge in such a way that higher quality decisions are made by the group than could be done by the

group's best member [14]. Positive group gains such as these effects generally fall under the umbrella term of "synergy", operationally defined as performance growth greater than $1 \times N$ in the number of group members, N.

A list of simple human-robot team organizations was compiled by Yanco and Drury [15]. This list includes (a) the conventional one-to-one relationship between human and robot, (b) teams of coordinating robots, such as arise in swarm robotics [16], [17] and "playbook-style" supervisory control [18], and (c) teams of coordinating humans managing one or more robots, such as those that have proven useful in search and rescue [6]. Yanco and Drury [15] intentionally left out the condition in which multiple operators controlled multiple robots without explicit coordination on either side since they felt that some explicit coordination was required, but "call center" metaphors have been proposed for MOSC with minimal coordination [19], [20].

III. METHODS

To explore constraints on MOSC robot control, we conducted a 2x2 experiment using novice operators.

Experiment Conditions	
One Operator Bombs-Only	Two Operators Bombs-Only
One Operator Bombs-and-Mines	Two Operators Bombs-and-Mines

TABLE I
EXPERIMENT CONDITIONS

A. Experiment Design

The first independent variable was controlling (a) a homogeneous team of 6 bomb-sniffer robots or (b) a heterogeneous team of 6 bomb-sniffer and 8 mine-sweeper robots. The second independent was working (a) alone or (b) with a teammate. A total of 48 (28 male, 20 female) paid subjects participated in the study: 16 of the participants worked alone, and 32 were part of a team (making 16 teams). Each subject participated in only one condition, making this a between-subjects design. Ages ranged from 19 to 27, with a median age of 22. Of the 16 teams, 9 were married couples and the remaining 7 were friends who signed up together (3 were all-male, 2 were all-female, and 2 were male/female pairings). Participants were told that the experiment would take no more than 75 minutes even if all tasks were not completed during that time.

Each subject completed a short, self-paced training course (20-30 minutes) consisting of a series of on-screen videos which explained the tasks, the interface, and the controls they would use to complete the tasks. All participants were given opportunities to practice what was explained, and teams received an additional practice of three to four minutes to experience how communication would be handled. The additional training for teams is a potential confounding factor, but statistical analysis of the results suggests that the additional training time did not have a significant impact on performance.

The experiment simulated an exercise in which subjects took on the role of explosive ordinance disposal (EOD) workers. Subjects had to search through a set of buildings searching for explosive devices. "Bomb-sniffing" robots searched for bombs inside of buildings by following the "chemical plume" surrounding the bomb-sniffers. "Mine-sweeping" robots were used to search for mines buried outdoors by sweeping an operator-designated search area.

Due to the different way in which the "bomb-sniffing" and "mine-sweeping" robots are operated, they constitute two logically separate tasks, with a limited amount of overlap. The mine-sweeping robots generally require much less input since they are automated while searching for mines, whereas the bomb-sniffers require significant operator control since there is no automated behavior for search. Mines were cleared by having a mine-sweeper pass over them while in search mode (a mode initiated by the user). Bomb-sniffers were cleared by a user placing an icon on top of the bombs. In order to find the bomb's location, the operator had to follow a colored gradient.

Teams of two operators could see the same virtual workspace, as well as the results of each other's commands. Beyond this, there was no coordination enforced, but communication was allowed through a chat window. Robots acted independently of each other, and the operators did not have to agree upon commands to go to the robot, and in fact could issue opposing commands which were resolved by having the robot simply following the last command received.

B. Software Environment

1) Maps: Three maps were created: one for training and two for testing. Map-presentation was counterbalanced. Figure 1 shows one of the testing maps used from both the user interface view (only a satellite image) and the simulator view (shows floor plans plus positions of bombs and mines). Experiment results indicated that map type had very little impact on results.

2) Autonomy: Each robot had a virtual sensor which detected obstacles/walls in front of it. When a wall was detected, it was automatically added to a global map shared by all robots, and displayed in the interface as bright blue lines.

Each robot had the ability to follow paths set by the operator manually or to automatically generate their own path to a user-specified go-to point. Manually created paths were generally shorter and more direct than automatic paths, but when an obstacle was encountered the robot would get stuck until the operator specified a new path. Automatic paths were less efficient, but would automatically re-plan to avoid obstacles discovered en route. Users also had the option of manually controlling the robots through the use of arrow keys on the keyboard.

Additionally, mine-sweeping robots had a "search mode" in which the user could draw out an area to be searched for mines, and the robot would create a lawnmower search path that attempted to fit the area the user specified. In search mode, robots could get stuck on obstacles but it was necessary for operators to use this mode since mine-sweepers could
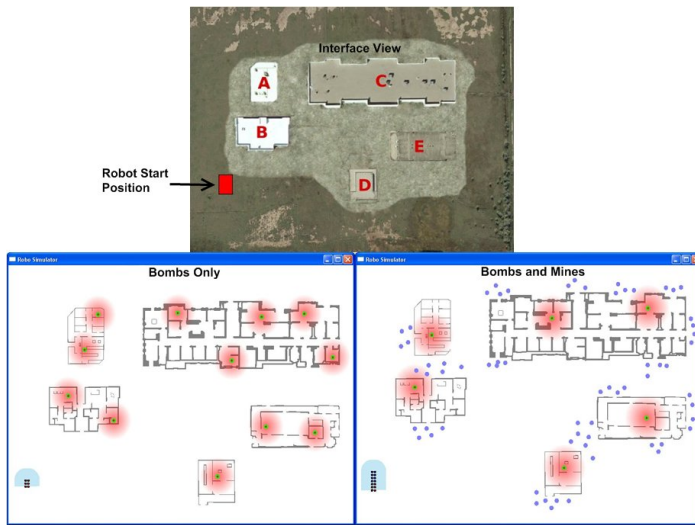
Fig. 1. One of the maps used in the experiment. The red areas represent the chemical plume put out by the bombs which is detected by the bomb-sniffing robots, the blue dots represent mines, and the blue area immediately around the robot represents their scanner range in which obstacles can be detected.

only automatically detect and disarm mines while in this mode. Although participants could not completely trust mine-sweepers to complete a search area without becoming stuck, participants still had had to rely on the autonomy to detect and disarm mines.

Bomb-sniffing robots had no search mode which would auto-detect bombs; they only indicated how close they were to a bomb using a color gradient. This meant that operators had to pay much closer attention to robots near a bomb-sniffer, and most chose to drive them manually when they began to see the gradient displayed near a bomb-sniffer.

3) Secondary Task: We created a secondary task in which participants had to respond to a "commander" who would ask them questions about the environment. These questions were picked at random from a predefined list, and would appear in a chat window located at the bottom left of the interface every 12-13 seconds. If a question was not answered, it would be repeated instead of asking a new question. For every commander question, 4-5 meaningless distractor messages would appear in a different color. Chat messages from a teammate were displayed in color different from channel chatter and commander questions.

4) Scoring: Participants were shown a running score that was based on their performance. Positive points were awarded for disarming bombs and mines, and points were taken away for losing a robot due to an explosion, incorrectly placing an icon to mark a bomb, and as a factor of time (one point per second). The score was not affected by participants' answers to questions from the "commander", though in hindsight this may have been useful.

C. Dependent Variables

Task performance included the number of bombs found, the number of mines found, the amount of time robots were

utilized, and several others. Workload was measured indirectly by measuring subjects' performance on a secondary task (answering questions from the commander), and by a subjective survey based on NASA TLX [21]. Fan-out was measured directly by counting the number of robots actively being controlled by each individual at 30 second intervals, and then averaged over the first 20 minutes of experiment time.

IV. RESULTS

For operator teams, we estimate individual performance by dividing team results by two, thus allowing comparisons with individuals working alone. This sets a standard such that an individual working as part of a team must show statistically significant improvements over individuals working alone.

We limit our examination to the first 20 minutes of each experimental run. For statistical analysis of means, we used the results recorded at the 20 minute mark, and then corroborated those results by looking at trends over the entire 20 minute time period. All significance values reported were calculated using an independent two-tailed T-test.

In this paper, we summarize only those results relevant to MOSC performance.

A. Primary Task Performance

In this section we present primary task performance.

1) Bombs Only: In the bombs-only condition, both individual operators and operator teams had twelve bombs distributed across five buildings. Figure 2 shows the cumulative total of
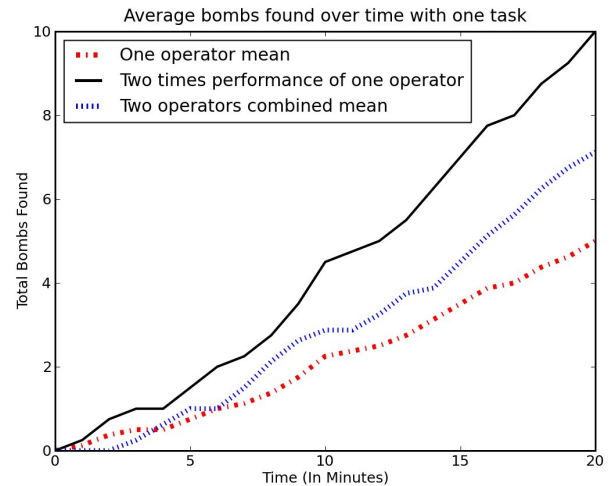

Fig. 2. Cumulative total of bombs found over a 20 minute period in the bombs-only condition.

bombs found over a 20 minute period. Eight operators working alone found an average of 5 bombs, while operators working in teams found an average of 3.56 bombs per operator. The difference was not statistically significant (p=0.192). These results suggest no more than $2\times$ growth for primary task performance using teams of operators.

We believe that this represents what Olsen calls *task saturation*. Task saturation occurs when adding robots does not

increase performance. For the bomb-sniffing task, there are only 6 bomb-sniffer robots and 5 buildings. Although the two operators controlled more robots than a single operator (2.19 robots per operator for the team vs. 3.47 robots for the single operator, p=0.0004), the performance of the two operators is not more than twice the performance of the single operators because adding extra robots doesn't make the task go much faster in the 5 buildings total.

2) Bombs and Mines: In the bombs and mines condition, operators had 6 bomb-sniffer robots and 8 mine-sweeper robots at their disposal to complete this task. This meant that there were 14 robots to monitor and command compared to only 6 in the bombs-only condition. Operators searched 5 buildings to find 6 bombs, and searched the area around those buildings to try to find 65 hidden mines. Based on subjective and objective workload evaluations, this condition was significantly more taxing than the bombs-only condition.
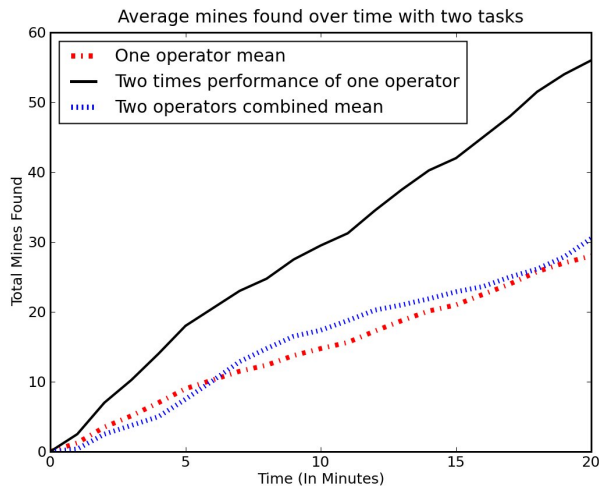

Fig. 3. Cumulative total of mines found over a 20 minute period in the bombs-and-mines condition.

The average number of mines fell from 28 per individual working alone to 15.25 per individual working on a team; see Figure 3. We propose that this is caused by something that we will call *task diffusion* which indicates that the task gets harder and less relevant the longer operators work on it. Most mines were primarily clustered around entrances to the buildings (which the subjects were told), with the remainder scattered throughout the environment. The mines near the entrances were relevant initially since they needed to be disarmed before bomb-sniffers could begin their work, but as these mines were disarmed the relevance of the mines decreased and the difficulty of finding the remaining mines increased.

B. Workload and Fan-Out

For the bombs-and-mines condition, individuals on teams answered 42% more questions from the "commander" than operators working alone (p=0.042) and Figure 4). Additionally, reported mental activity and mental work fell by 36% (p=0.0009) and 20% (p= 0.052)respectively, and time pressure

felt was reduced by 27% (p=0.0126). Additionally, 75% of subjects working alone reported that they were either "Unsuccessful" or "Somewhat Unsuccessful" in accomplishing the tasks while nearly 69% of subjects working together reported at least "Satisfactory" success.
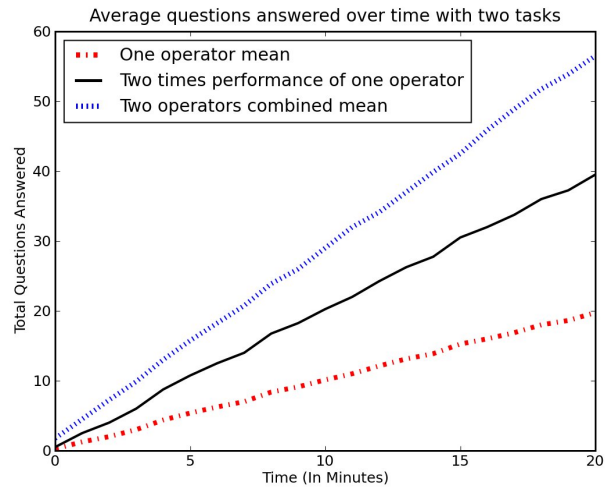

Fig. 4. Cumulative total of questions answered in the bombs-and-mines condition.

For both the bombs-only (p=0.0004) and the bombs-and-mines condition (p=0.015), fan-out declines by more than 35% in the number of robots controlled by individuals that worked in teams compared to those who worked alone . Since performance did not decline, this implies that a decrease in fan-out does not always correspond to a decrease in performance. It is not apparent from the data whether the lower fan-out contributed to lower workload for operator teams, but we suspect this is the case. If this is true, team members had additional capacity and were limited by the constraints the experiment imposed (task diffusion and task saturation). Since task saturation and task diffusion did not allow this spare capacity to be used in this experiment, we will use agent-based simulation to better understand how spare capacity might be useful for MOSC teams.

C. Analysis

In this section, we identify some elements of team behaviors that contributed to high performance.

1) Specialization: By having two different types of tasks, searching for bombs and searching for mines, it was possible that teams would have members who specialized into performing one task or the other. Test subjects were instructed that they would be more effective if they cooperated, but were not assigned a specific role. Within the operator teams, it was common (but not universal) for one member to focus on searching for mines, while the other team member focused on searching inside of buildings for bombs. If we look at the number of mines found by individuals in a team, and then separate the operators into two groups based on which operator found more mines than their teammate, the average number of

mines found by the "specialist" teammates is 26.13, compared to an average of 4.38 for the "non-specialist" teammates. This number is much closer to the number of mines found by operators working alone, and is consistent with the explanation that there was a performance "cap" for mine-sweeping robots due to the experimental setup.

To evaluate specialization, we correlated combined performance data for teams with the amount of specialization that took place within the team. We calculated a "specialization score" using the following formula:

Specialization = | Teammate(A) Fan-out − Teammate(B) Fan-out |

A higher value indicates that there was more specialization among the team (since there is a greater disparity in the number of robots of one type they controlled), and hence a greater division of labor occurring.

Given these measurements, we found no correlation (coefficient = -.04) between bomb specialization and the number of bombs teams found. There was, however, a strong positive correlation (coefficient = .89) between mine specialization and the number of mines teams found (see Figure 5). This suggests that specialization is a key element of effective MOSC performance for heterogeneous robot teams.
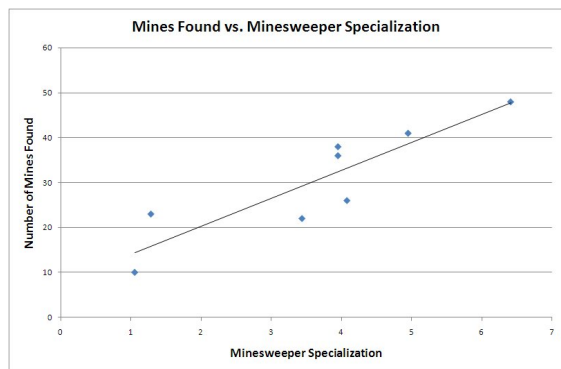

Fig. 5. Scatter plot correlating mines found with mine-sweeper specialization with a correlation coefficient of .89.

V. SPARE CAPACITY AND AGENT-BASED SIMULATION

The evaluation of the human factors experiments suggests that when humans worked together they had lower workload and more capacity for performing secondary tasks. In this section, we use agent-based simulation to evaluate how spare capacity can impact MOSC performance.

A. Modeling

Since the goal of this section is to identify the performance potential created by spare capacity, it is necessary to model two types of agents: robots and humans. We begin by addressing modeling robots.

1) Modeling Robots: Neglect and Interaction Time: The interaction characteristics of a single robot have been modeled in prior work as an ordered pair of (Neglect-Time,InteractionTime). Neglect time (NT) represents the amount of time that a robot can operate autonomously without human intervention before the robot's average performance drops below a *minimum performance threshold*. Interaction time (IT) represents the average amount of time required for an operator to bring the robot from the lowthreshold to an *upper performance threshold*, what might be called the delegation threshold where the human is satisfied that the robot can safely be neglected again.

In practice, computing NT and IT are very noisy [22]. However, in an agent-based simulation it is useful to abstractly represent NT and IT as idealized quantities; future work will model NT and IT using statistical models. We choose to model NT and IT in this paper as percentages of total time. For example, suppose that a human interacts with a robot for 10 hours in such a way that interaction begins at the minimum performance threshold and neglect begins at the upper performance threshold. If the human neglects the robot for 9 hours and interacts for 1 hour then NT=0.9 and IT=0.1.

In practice, high NT and low IT produce high fan-out, since Olsen's fan-out model gives

$$FO = 1 + \frac{NT}{IT}.$$

Since we represent NT and IT as a percentage of total time,

$$NT + IT = 1.$$

It is useful to make a simplifying assumption that the level of autonomy is given by

$$LOA = NT$$

so high levels of autonomy correspond to high neglect times. Since we assume that $NT+IT = 1$ for all LOAs, we implicitly restrict the set of possible (NT, IT) pairs to a small subset, that is $LOA \in (0, 1)$. This subset is important because (a) all LOAs within this set satisfy payoff dominance, meaning that no LOA within the set is dominated by another possible LOA that has both higher NT and lower IT, and (b) fan-out can be expressed very simply by observing that $NT = 1 - IT$ yielding:

$$FO = 1 + \frac{NT}{IT} = 1 + \frac{1 - IT}{IT} = \frac{1}{IT},$$

that is, fan-out grows as interaction time shrinks.

This model makes a number of simplifying assumptions, so it is useful to clearly state these assumptions, identify the limitations implied by the assumptions, and discuss whether these limitations are ecologically valid for real human-robot interaction. The first assumption is that $NT + IT = 1$. This assumption implies that the operator perfectly follows the "two-threshold model", meaning that the operator always begins to neglect the robot when its performance reaches the upper threshold and engages the robot when its performance reaches the lower threshold. The operator never opts to interact with the robot for a little extra time to raise its performance a little above threshold, and the operator is never distracted by other tasks that would cause him or her to neglect the robot a little longer than needed. This is a severe limitation since

operators will rarely conform to such ideal behavior. In terms of what the agent-based simulation can tell us, however, this limitation allows us to systematically study what might happen if operators can deviate from this ideal, average-case model. Indeed, we will shortly show that a little extra interaction time can make a dramatic difference in performance.

The second assumption is that all levels of autonomy satisfy $NT + IT = 1$. This assumption implies a type of *conservation of interaction and neglect time* across LOAs, meaning that whenever one LOA has higher neglect tolerance its corresponding interaction efficiency is less effective. To understand this assumption and its consequences, it is useful to discuss how performance changes with neglect and interaction. We do this in the next section.

2) Modeling Interaction: Performance: Associated with each LOA is a performance and cost. When robots are doing productive work then they are contributing to the mission of the team. However, when robots require interaction from a human manager, then they consume a limited shared resource the managers attention. In prior work, the performance of the robot as a function of neglect and interaction have been called the *neglect impact* and *interface efficiency* curves, respectively [22]. Costs have been implicitly discussed in prior work as a constraint satisfaction problem wherein a human has a fixed budget of time that must be scheduled across multiple robots. In this section, we explicitly model performance during periods of interaction and neglect subject to a constrained budget of human attention.

We begin by modeling performance during periods of interaction, the so called interface efficiency curve which would be better called the *interaction impact curve* in the context of this paper. The most simple model of the impact of interaction on a robot's performance is to assume that instantaneous performance increases linearly, yielding

$$P_{\text{interaction}}(t) = m_I t \qquad (1)$$

meaning that a robot's instantaneous performance increases linearly while a human interacts with it. Given that we are modeling the ideal, average operator who always begins to neglect the robot at the upper performance threshold, it is reasonable to assume that performance grows linearly as long as the operator follows ideal behavior. For real operators who may interact a bit longer, performance can continue to grow until interaction no longer improves performance. We will model this by using unitless interaction impact curves with a minimum instantaneous performance set to zero and a maximum instantaneous performance set to unity, meaning, maximum performance. This latter assumption simply causes performance to saturate at some point if the operator interacts beyond the nominal value for IT. Note that actual performance is the area under the curve, that is, the integral of instantaneous performance.

The instantaneous performance during periods of interaction is illustrated in Figure 6. In the figure, performance grows linearly while the operator interacts with the robot. When the performance reaches the upper threshold after $t = IT$ seconds
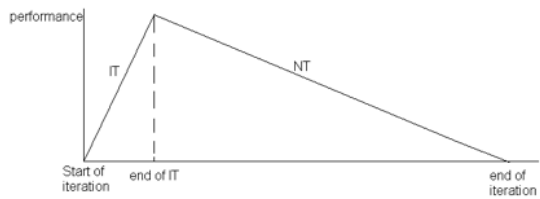

Fig. 6. Ideal interaction and neglect impact curves. Performance increases linearly during interaction and decreases linearly during neglect.

have passed, performance is given by $P_{\text{interaction}}(t = IT) = m_I IT$. At the end of interaction, performance drops to the lower threshold. Without loss of generality, we set the lower threshold to zero. In Figure 6 an *iteration* is the amount of time that elapses over one cycle of interaction and neglect, assuming that interaction terminates at the upper threshold and resumes at the lower threshold. Since $NT + IT = 1$, one iteration ends when time reaches one. Given this constraint, we can compute the slope of the performance curve during neglect as $m_N = -m_I \frac{IT}{NT}$ yielding

$$\begin{aligned} P_{\text{neglect}}(t) &= m_N t + c \qquad (2) \\ &= -m_I \frac{IT}{NT} t + c \end{aligned}$$

3) Free Capacity for Homogeneous Teams: In this section, we use agent-based simulation to explore how operator free capacity (OFC) can benefit a team of homogeneous robots managed by a single operator. This requires us to integrate OFC into the agent-based model. Let R denote the total number of robots available. At maximum human capacity, $R = FO$; this means that the largest number of usable robots on a homogeneous team is equal to the fan-out. Spare capacity occurs when humans operate at less than maximum capacity, yielding fewer robots being used on a team. Let R_{eff} be the effectual number of robots being used when a human does not work at full capacity. Operator free capacity can then be defined as

$$OFC = 1 - \frac{R_{\text{eff}}}{FO}, \qquad (3)$$

that is, OFC is the proportion of an operator's capacity that is not being used to manage robots. Since the team is homogeneous, all robots on the team have the same average IT so OFC can be interpreted in temporal terms as the percentage of the operator's time that is not being used to manage the robots.

The temporal interpretation of OFC emphasizes the *discretionary* nature of spare capacity; operators may or may not choose to use this free time to service robots. In Figure 7, the operator interacts with the robot as usual during phase A and then neglects the robot during phase B. Rather than allowing the instantaneous performance to reach the lower threshold, if the operator has spare capacity he or she can discretionarily choose to interact with the robot again before it reaches the lower performance threshold as in phase C. The area under
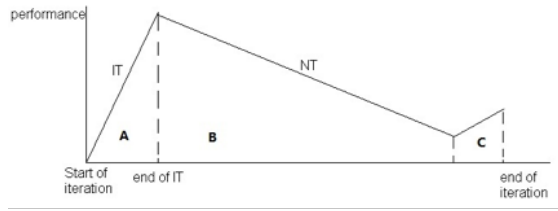
Fig. 7. If the operator is not at maximum capacity, he or she can interact with the robot before it reaches the lower neglect threshold.

the curve, the accumulated performance over time, exceeds performance if the operator had not had spare capacity.

Since the operator need not interact with the robot during free time, it is useful to model how often the operator uses free time to improve robot performance. Although the operator may have other things to do during free time, we will refer to the likelihood that the free time is not used to service robots as the *lazy factor* and denote it ℓ. This factor is interpreted as follows: when $\ell = x$ this means that the operator only interacts with the robot during free time $x\%$ of the time.
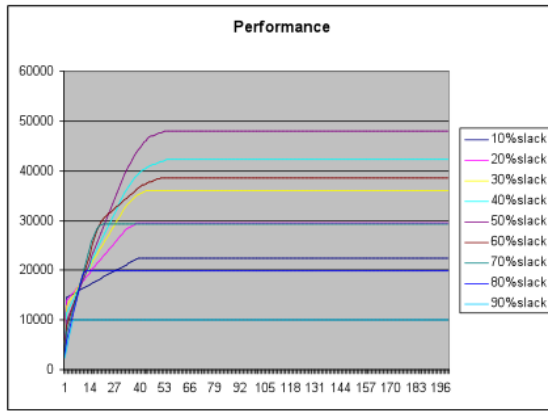

Fig. 8. For $\ell = 0.1$, results of performance for various OFC levels as a function of simulation time. $LOA = 0.9, m_I = 1$. The term *slack* is the simulation term corresponding to OFC.

Given the parameters of the agent-bsed simulation given in the caption to Figure 8, a series of agent-based simulations was conducted. Figure 8 shows total performance as a function of OFC for $\ell = 0.1$. Notice that total team performance peaks when OFC is 50%. Although this particular value of OFC is a function of the specific model parameters chosen in the agent-based simulation, the result is robust over different values of the lazy factor. Since instantaneous performance of all agents saturates at unity, the plots in Figure 8 reach a steady-state value.

When we vary the lazy factor over the interval $(0.1, 0.9)$ (discretized in 0.1 increments) and measure performance, we can take the maximum value of various slack times. The results are shown in Figure 9 and indicate that the having 50% OFC produces higher team performance. Simply put, the ability to use slack time to operate above the performance thresholds has the potential for improving team performance.
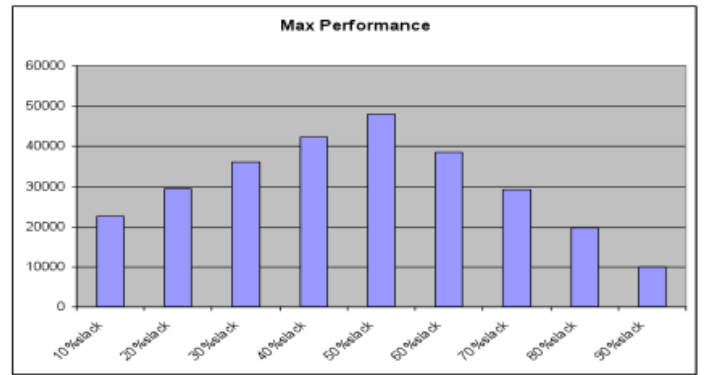

Fig. 9. Results of performance as a function of OFC, $max_{\ell \in (0.1, 0.9)} \sum_{\text{agents}} \sum_t P(OFC, t). LOA = 0.9, m_I = 1.$

B. Free Capacity for MOSC Teams

As shown in the previous section, when an operator is given 50% slack time, a single operator has the potential for managing multiple robots at higher performance than they would achieve if they maximized fan-out. To apply this in a MOSC context, what kind of performance gain or benefit would there be if instead of a (maximally loaded) single operator we had two (less-loaded) operators each controlling an independent team?

To do this in a useful agent-based simulation, it is helpful to evaluate performance when there is a cost of communication between the operators. Communication cost will reduce the total time the operators have to interact with the robots meaning performance will be lower. We model this as a multiplicative factor to the final performance of the two teams, meaning that the performance of both operators suffers the same amount due to communication cost. Let $o \in (0, 1)$ denote communication overhead so that team performance is discounted by $(1 - o)$.
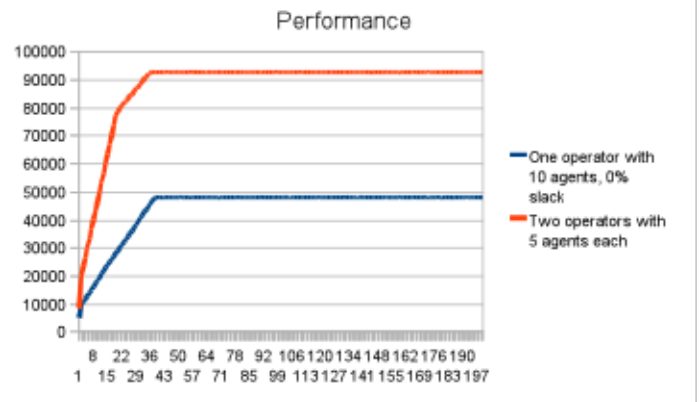

Fig. 10. Results of team performance versus individual performance: $LOA = 0.9, m_I = 1, o = 0.2.$

Figure 10 shows the results of a team of two operators versus a single operator. Performance is shown for two operators who (a) could potentially control 8 agents each, (b) have enough free capacity that they actually control 5 agents each

agents, and (c) have a communication overhead of $o = 0.2$. This is shown in contrast to a single operator controlling a the maximum number of 10 robots without any free capacity. Thus, two operators can be more effective when controlling the same number of robots as a single operator simply because spare capacity may allow them to improve performance even when there is significant communication overhead.

VI. CONCLUSIONS

In this paper, we identified two task constraints that may limit performance of MOSC teams: task saturation and task diffusion. These two constraints imply that increasing fan-out by increasing the number of operators may not produce higher performance if the task does not allow performance to increase as fan-out increases. The experiments also provides evidence that specialization among operators across different tasks can potentially improve performance, and that sharing the responsibility for managing multiple robots may provide spare capacity.

Using agent-based simulation, we then explored how spare capacity could potentially improve performance of a single operator. If a single operator has spare capacity then he or she can discretionarily choose to improve individual robot instantaneous performance and thereby improve overall team performance. This potential benefit from spare capacity even holds when operators do not use all of there time to manage robots; spare capacity is useful even if operators do not use their discretionary time effectively.

This discretionary time can have a strong potential impact on the potential usefulness of a team of robots when managed by multiple operators. An agent-based simulation indicated that two operators who have spare capacity, capacity that stems from managing fewer robots than they would if managing robots alone, can outperform a single operator using more robots. This result holds even when communication cost between operators is non-trivial.

There are important limitations to this work, including the precise modeling assumptions in the agent-based simulation. Results on the optimal amount of spare capacity are a function of these modeling assumptions. However, the potential usefulness of this spare capacity is something that deserves future work. Future work should also include adding to the agent based simulation errors, switch costs, and other human factors parameters. Similarly, task parameters and constraints, including saturation and diffusion, should be evaluated in more sophisticated agent-based simulations, and the resulting observations should be used as hypotheses for a series of follow-on human factors work.

ACKNOWLEDGEMENTS

REFERENCES

[1] D. R. Olsen and S. B. Wood, "Fan-out: Measuring human control of multiple robots," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. Vienna, Austria: ACM, 2004, pp. 231–238.

[2] P. A. Hancock, M. Mouloua, R. Gilson, J. Szalma, and T. Oron-Gilad, "Is the UAV control ratio the right question?" *Ergonomics in Design*, vol. 15, no. 1, pp. 7–30, 2007.

[3] M. A. Goodrich and D. R. Olsen, "Seven principles of efficient human robot interaction," in *IEEE International Conference on Systems, Man and Cybernetics*, vol. 4, 2003, pp. 3942–3948 vol.4.

[4] M. L. Cummings, C. E. Nehme, J. W. Crandall, and P. Mitchell, "Predicting operator capacity for supervisory control of multiple UAVs," *Studies in Computational Intelligence (SCI)*, vol. 70, pp. 11–37, 2007.

[5] M. A. Goodrich, M. Quigley, and K. Cosenzo, "Task switching and multi-robot teams," in *Proceedings from the 2003 International Workshop on Multi-Robot Systems*, L. E. Parker, F. E. Schneider, and A. C. Shultz, Eds., vol. 3. Springer Netherlands, 2005, pp. 185–196.

[6] R. R. Murphy and J. L. Burke, "Up from the rubble: Lessons learned about HRI from search and rescue," in *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 2005, pp. 437–441.

[7] M. Cummings, S. Bruni, S. Mercier, and P. Mitchell, "Automation architecture for single operator, multiple UAV command and control," *The International C2 Journal*, vol. 1, no. 2, pp. 1–24, 2007.

[8] I. D. Steiner, *Group Process and Productivity*. New York, NY: Academic Press, 1972.

[9] B. A. Nijstad, W. Stroebe, and H. F. M. Lodewijkx, "Production blocking and idea generation: Does blocking interfere with cognitive processes?" *Journal of Experimental Social Psychology*, vol. 39, no. 6, pp. 531–548, 2003.

[10] L. A. Penner and J. P. Craiger, "The weakest link: The performance of individual team members," in *Teams: Their training and performance*, R. W. Swezey and E. Salas, Eds. Ablex Publishing Corporation, 1992.

[11] A. Pinsonneault, H. Barki, R. B. Gallupe, and N. Hoppen, "Electronic brainstorming: The illusion of productivity," *Information Systems Research*, vol. 10, no. 2, pp. 110–133, 1999.

[12] J. Macmillan, E. E. Entine, and D. Serfaty, "Communication overhead: The hidden cost of team cognition," in *Team Cognition*, E. Salas and S. M. Fiore, Eds. American Psychological Association, 2004.

[13] R. B. Zajonc, "Social facilitation," *Science*, vol. 149, no. 3681, pp. 269–274, 1965.

[14] B. E. Collins and H. S. Guetzkow, *A Social Psychology of Group Processes for Decision-Making*. Wiley, 1964.

[15] H. A. Yanco and J. L. Drury, "Classifying human-robot interaction: An updated taxonomy," in *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, vol. 3, 2004, pp. 2841–2846.

[16] T. Balch and L. E. Parker, Eds., *Robot Teams: From diversity to polymorphism*. A. K. Peters Ltd, 2002.

[17] K. Lerman, A. Galstyan, A. Martinoli, and A. Ijspeert, "A macroscopic analytical model of collaboration in distributed robotic systems," *Artificial Life*, vol. 7, no. 4, pp. 375–393, 2002.

[18] R. Parasuraman, S. Galster, P. Squire, H. Furukawa, and C. Miller, "A flexible delegation-type interface enhances system performance in human supervision of multiple robots: Empirical studies with Roboflag," *IEEE transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, vol. 35, pp. 481–492, 2005.

[19] J. V. Nickerson and S. S. Skiena, "Attention and communication: Decision scenarios for teleoperating robots," in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 2005, p. 295.

[20] M. Lewis, J. Wang, and P. Scerri, "Teamwork coordination for realistically complex multi-robot systems," in *NATO Symposium on Human Factors of Uninhabited Military Vehicles as Force Multipliers*, 2006.

[21] S. Hart and L. Staveland, "Development of NASA-TLX (Task load index): Results of empirical and theoreticalresearch." in *Human mental workload*, P. Hancock and N. Meshkati, Eds. Amsterdam: North Holland Press, 1988, pp. 139–183.

[22] J. W. Crandall, M. A. Goodrich, D. R. O. Jr., and C. W. Nielsen, "Validating human-robot interaction schemes in multi-tasking environments," *IEEE Transactions on Systems, Man and Cybernetics — Part A: Systems and Humans*, vol. 35, no. 4, 2005.