A Flexible Delegation-Type Interface Enhances System Performance in Human Supervision of Multiple Robots: Empirical Studies With RoboFlag

Raja Parasuraman, Scott Galster, Peter Squire, Hiroshi Furukawa, and Christopher Miller

***Abstract*—Three experiments and a computational analysis were conducted to investigate the effects of a delegation-type interface on human supervision of simulated multiple unmanned vehicles. Participants supervised up to eight robots using automated behaviors ("plays"), manual (waypoint) control, or both to capture the flag of an opponent with an equal number of robots, using a simple form of a delegation-type interface, Playbook. Experiment 1 showed that the delegation interface increased mission success rate and reduced mission completion time when the opponent "posture" was unpredictably offensive or defensive. Experiment 2 showed that performance was superior when operators could flexibly use both automated behaviors and manual control, although there was a small increase in subjective workload. Experiment 3 investigated additional dimensions of flexibility by comparing delegation interfaces to restricted interfaces. Eight interfaces were tested, varying in the level of abstraction at which robot behavior could be tasked and the level of aggregation (single or multiple robots) to which plays could be assigned. Performance was superior with flexible interfaces for four robots, but this benefit was eliminated when eight robots had to be supervised. Finally, a computational analysis using task-network modeling and Monte Carlo simulation gave results that closely paralleled the empirical data on changes in workload across interface type. The results provide initial empirical evidence for the efficacy of delegation-type interfaces in human supervision of a team of multiple autonomous robots.**

***Index Terms*—Automation, delegation, human–robot interaction, Playbook, unmanned vehicles.**

I. Introduction

ROBOTS and unmanned vehicles (UVs) with increasingly sophisticated capabilities are being developed for use in many aerial, ground, surface, and underwater environments. Such robots can move and navigate autonomously, engage in goal-directed behaviors, and communicate with and provide feedback to human supervisors [1]. Human supervision is necessary to manage unexpected events and to ensure that mission goals are met [2]. Close attention must therefore be paid to the design of the human–robot interface, so as to allow for effective teaming, communication, and mission success. However, interfaces for human–robot interaction are still largely being developed ad hoc, based more on technological capabilities than on mission requirements and the needs of the human users [3]–[5].

Previous work in human–computer interaction can inform the design of human–robot interfaces. Nevertheless, autonomous robots are sufficiently different from most computer systems as to require new research and design principles [6], [7]. Design approaches should be based on empirical and computational modeling studies [8] rather than only on theoretical assertions, technology demonstrations, or field evaluations of prototype systems (although field studies are important in later stages of development). We provide empirical evidence in this article from three experiments that examined the effects of a delegation-type interface [9], [10] on human performance in supervising a team of multiple simulated robots.

Supervision of multiple robots is currently highly labor intensive. For example, present day military UVs typically require several operators for each single vehicle employed. Consequently, a goal for developers has been to reduce the operator to UV ratio, so that a small group of human operators (M) can control a much larger number (N) of UVs that are more autonomous and work cooperatively in teams [11]–[14]. However, while reducing the $M:N$ ratio may be a useful engineering goal, it should be seen only as one potential enabler of overall mission efficiency, rather than an end in itself. A more basic issue is what types of operator interfaces lead to a greater probability of mission success for human–robot teams [15].

Studies of coordination between humans and automated agents has revealed both benefits and costs of particular interfaces [16], [17]. Overreliance, reduced situation awareness, uncalibrated trust, mode errors, loss of operator skill, and unbalanced mental workload are among the costs that have been found to be associated with particular human–automation interfaces [18], [19]. The costs are not inevitable, but are a consequence of particular designs, and can be mitigated by interfaces that provide feedback to the operator on automation states [8], [16], communicate with humans in ways that follow the norms of human-human communication [20], or support human information-gathering activities [19]. Systems that minimize human participation in higher level decision-making processes by providing automated solutions can enhance overall system efficiency and reduce operator mental workload,

Manuscript received July 31, 2004; revised February 22, 2005. This work was supported in part by the Defense Advance Research Projects Agency (DARPA) under MICA Program through Contract F30602-01-2-0577 (S. Heise, monitor) and in part by the Army Research Laboratory under Contract 6004.005.01-C (M. Strub, monitor) to R. Parasuraman. The Playbook (trademark, SIFT, Inc.) work was funded by a Phase II SBIR grant from DARPA IXO to C. Miller from the U.S. Army Aeromedical Center under Contract DAAH01-03-C-R177. This paper was recommended by Associate Editors Julie A. Adams and Marjorie Skubic.

R. Parasuraman and P. Squire are with the Arch Laboratory, George Mason University, Fairfax, VA 22030 USA.

S. Galster is with the Air Force Research Lab, Wright-Patterson AFB, OH 45433 USA.

H. Furukawa is with Tsukuba University, 305-8577 Tsukuba, Japan.

C. Miller is with SIFT Inc., Minneapolis, MN 55401 USA.

Digital Object Identifier 10.1109/TSMCA.2005.850598

but only if the automation is completely reliable, a requirement that can seldom be achieved in practice [21]–[24]. Furthermore, even if full automation of decision-making functions can be made more reliable, the required computational and engineering efforts may be considerable, with only diminishing returns in benefits obtained. This was illustrated in a recent study in which the tradeoff between the cost of developing higher levels of automation and the resultant gain in efficiency was examined using human telerobotic control of a simulated "sheepdog" herding a group of simulated "sheep" [25].

These considerations suggest that the interface between humans and technology should be adaptive or adaptable, rather than fixed or static [26]–[30]. In adaptive systems, decisions to invoke automation or to change the degree of autonomy are made by the system [28]. In contrast, in adaptable systems, such changes are left to the user [30], [31]. The performance benefit of adaptive compared to static automation is well documented [30], [32]. However, because human operators may be unwilling to accede to the authority of a computer system that mandates when automation is to be used, user-adaptable automation has also been considered. Billings and Woods [33] cautioned that adaptive systems can be problematic if their behavior is unpredictable to the user. If the user explicitly invoked the automation, then presumably the unpredictability will be lessened, assuming that automation states and behaviors are made transparent to the user [8], [17]. But involving the human operator in making decisions about when and what to automate can increase mental workload. Thus, there is a tradeoff between increased unpredictability versus increased workload in systems where automation is invoked by the system or by the user, respectively [9], [10].

Delegation-type interfaces can allow adaptable automation to be implemented at a flexible and variable balance point in this tradeoff space. Humans should be able to delegate tasks to automation at times of their own choosing, and receive feedback on their performance. Delegation in this sense is identical to that which occurs in successful human teams. Delegation interfaces represent a particular type of real-time supervisory control [34] wherein the human sets an objective, provides more or less detailed instructions, and then tasks the automation to determine the best method to proceed toward the goal within the delegation instructions. Delegation architectures provide highly flexible methods for the human supervisor to declare goals and provide instructions and thereby choose how much or how little autonomy to delegate to automation on an instance-by-instance basis. An example of such a delegation architecture is the Playbook, which has been described elsewhere [9], [10]—so named because it is based on the metaphor of a sports team's book of approved plays and the selection of these plays by the team leader (e.g., the quarterback in American football) and executed by the team members (the other players).

The Playbook uses a hierarchical task model to provide a common language for a human supervisor to communicate goals and intents and a Hierarchical Task Network planning system [35] to understand, reason over, and either critique or complete partial plans provided by the human. This form of delegation interface permits the operator to "task" automation (such as robots or UVs): 1) at several functional levels of abstraction [36]; 2) by providing goals, plans, and/or constraints in any combination; and 3) by providing temporal, sequential, or conditional constraints on task performance at varying levels of depth.

Evaluations of different interfaces for human supervision of multiple robots can be informed by studies employing objective measures of human performance and strategy use. Such studies are still relatively rare [4], [26], [37]–[39]. In the present studies, we examined the use of simplified forms of the Playbook interface with the RoboFlag simulation environment [2], [4], [40]. A delegation interface gave the user the ability to task simulated robots, individually or in groups, at a minimum of two levels of abstraction: by providing designated waypoints for robot travel or by commanding higher level behaviors (or "plays"). The RoboFlag simulation was modified to emulate a typical UV mission involving a single operator managing a team of robots. The mission goal was to send the robots from a home area into opponent territory, access and obtain a specified target (the flag), and return home as quickly as possible with minimum loss of assets.

II. Experiment 1

One of the postulated benefits of delegation-type interfaces such as Playbook is that they allow for flexible use of automation in response to unexpected changes in task demands, while keeping the operator's mental workload in using the automation within a manageable range [9], [10]. In the first experiment we therefore explored the use of a delegation interface in a multiple UV simulation in which two sources of task demand were varied: 1) adversary "posture," in which the opponent engagement style was changed unpredictably between three types, offensive, defensive, or mixed; and 2) environmental observability, by varying the effective visual range of each robotic vehicle under the control of the operator. We hypothesized that the use of a simplified form of Playbook would allow users to respond effectively to unexpected changes in opponent posture and to increased uncertainty associated with limited visual range. We therefore assessed changes in how users tasked and supervised robots (using both manual control and the automation tools that were part of the delegation interface) as a function of these two factors. We also examined the effects of these factors on overall mission efficiency (success rate and time to completion) and subjective mental workload and situation awareness.

A. Methods

1) Participants: Eight males and ten females between the ages of 18 and 33 ($M = 23$ years) served as paid participants.

2) Apparatus and Procedure: RoboFlag consists of hardware robots developed to exhibit various forms of autonomous, cooperative behaviors [40] in games against another robot team such as soccer, "tag," and "capture the flag." The RoboFlag simulation, which was run on three separate personal computers (PCs) communicating under TCP/IP protocol, accurately captures the actions and states of the actual robots on computer displays. The human operator used one PC while another ran the opposing team and the third PC displayed a central processing executive (the "Arbiter") and collected the data. We modified the RoboFlag simulation to allow a single operator (blue team)

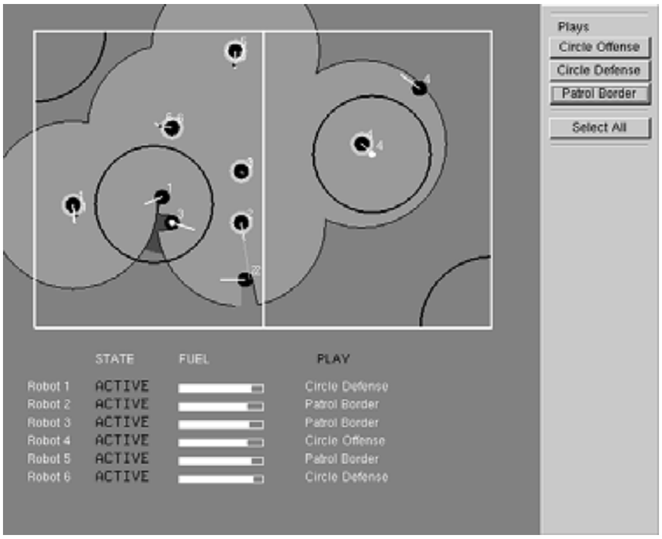


Fig. 1. Grayscale version of the RoboFlag simulation with simplified Playbook interface for blue team operator at a single moment of time during the mission. All six blue team robots (filled black circles with white rings) and four of the red team robots (filled black circles) within current visual range (merged gray circles) of the blue team are shown. Blue and red team flag areas are shown as open black circles on the left and right sides of the display, respectively, with the flags in each area being shown as filled white circles. Play selection buttons are on the top right-hand corner, and robot state and fuel status with corresponding assigned plays are at the bottom.

to compete against an automated opponent (red team) operating under scripted procedures that simulated different opponent "postures." The field of engagement was divided into two halves (see Fig. 1). The operator supervised six robots and had to send some or all of them into the opponent's field, capture their "flag" (located in at the center of a circular area in the opponent's half) and return safely to midfield, while simultaneously protecting their own flag (also located at the center of a circle in the home half). Each robot had a specified visual range of approximately 300° in the forward direction of movement, as represented by a circular ring of specified diameter around the robot (see Fig. 1). The operator could control robots either manually using point-and-click (waypoint) control, or through commanding "plays"—higher level aggregate behaviors. In this experiment, three plays were made available, circle offense, circle defense, and patrol border. In circle offense, a user-selectable number of blue team robots traversed to the red team's field, positioned themselves around the red team's flag circle, and attempted to capture the flag. In the circle defense play, the designated blue team robots circled around their own flag to prevent the red team from reaching and extracting their own flag. Finally, in patrol border user-selected robots positioned themselves along the dividing line in the field of engagement in order to defend their territory against incursions by the red team.

The opponent (red team) posture was varied to be offensive, defensive, or mixed. In the offensive condition, all six red team robots used *circle offense* to try to capture the blue team's flag and return it to the midfield line. When defensive, three red team robots defended the midfield line (*patrol border*) and the other three circled their own flag (*circle defense*) to prevent the blue team from reaching their flag. In the mixed condition three red team robots used *circle offense* and three *circle defense* (with none on *patrol border*). These three opponent postures were varied randomly and in an unpredictable (to the human operator of the blue team) way within each block of trials for a given robot visual range. The human operator could assign any number of the six blue team robots to an automated play or could control them manually by giving the robot a waypoint to move to (see Fig. 1).

Participants were trained by showing them the field of engagement, how to select and move robots, and use the automated plays. They were instructed that the only way a red team robot could be seen was if they entered into the visual range of the blue team robot, otherwise the red team robot was invisible to the blue team operator. They were shown the objective of capturing the opponent flag and crossing back into their own territory. Each participant completed a training trial for each of the nine experimental conditions. The objective in each condition was the same: cross into the opponent area with one or more robots, capture the opponent flag, and cross the midfield line while concurrently defending their own flag from capture by the opposing team. The mission was successfully completed when any one of the blue team robots returned to its half of the field with the red team flag. Participants were instructed to maximize the probability of mission success while minimizing mission completion time.

In a within-subjects design, three robot visual ranges (low, medium, high) were combined with three opponent postures (offensive, defensive, mixed). Visual Range was inversely related to observational uncertainty (e.g., low visual range=high uncertainty). Each participant completed five trials in each condition for a total of 45 trials. Visual Range was a block factor with Opponent Posture randomized within each block. Participants completed the National Aeronautics and Space Administration (NASA) Task Load Index (TLX) [41] and three-dimensional (3-D) Situation Awareness Rating Technique (SART) [42, pp. 3-1 and 3-17] after each of the three blocks. Objective performance measures included the percentage of missions successfully completed and mission completion time.

B. Results

1) Overall Performance: The performance data were submitted to a 3 × 3 analysis of variance (ANOVA) with factors of Visual Range (low, medium, high) and Opponent Posture (offensive, defensive, mixed). There was a significant effect of Opponent Posture on mission success rate, $F(2, 34) = 50.89, p < 0.01$ [see Fig. 2(a)]. Expectedly, the participants won 100% of the simulated missions when the opponent stance was purely defensive and no moves were made to capture the blue team flag. Excluding the defensive opponent posture condition, participants won significantly more games, $t(34) = 3.51, p < 0.05$, when they played against the red team on offense ($M = 75.6\%, SE = 2.6\%$) than when that team split the robots in the mixed condition ($M = 62.2\%, SE = 3.0\%$). Neither the effect of Visual Range nor the interaction between Visual Range and Opponent Posture were statistically significant for mission success rate.

A similar 3 × 3 ANOVA showed that mission completion times were significantly shorter when participants played against the red team offensive stance ($M = 36.48$ s, $SE = 0.47$ s) compared to the mixed posture ($M = 51.29$ s, $SE = 1.63$ s), $F(2, 34) = 177.92, p < 0.01$ [see Fig. 2(b)].

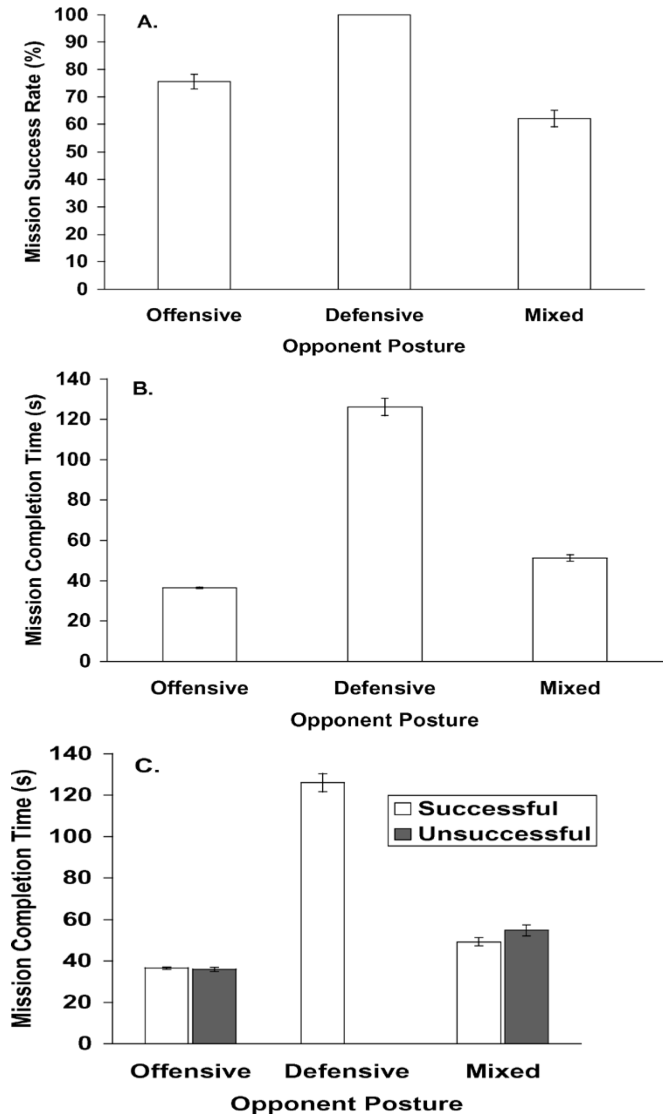


Fig. 2. Effects of opponent robot team posture (offensive, defensive, mixed) on (a) mission success rate, (b) mission completion time, and (c) completion time for successful and unsuccessful missions.

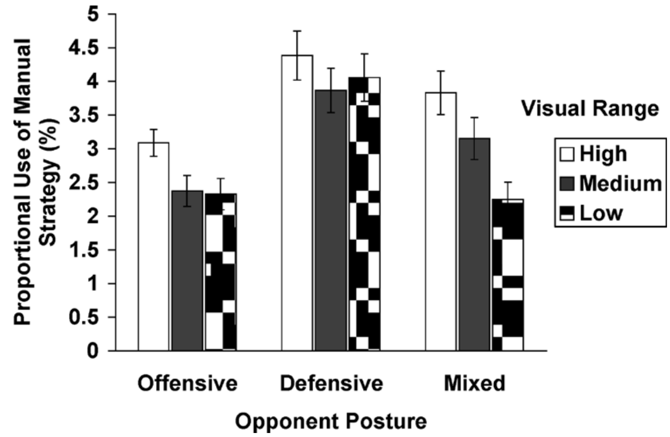


Fig. 3. Effects of opponent posture and robot visual range (low, medium, high) on proportional usage of manual control.

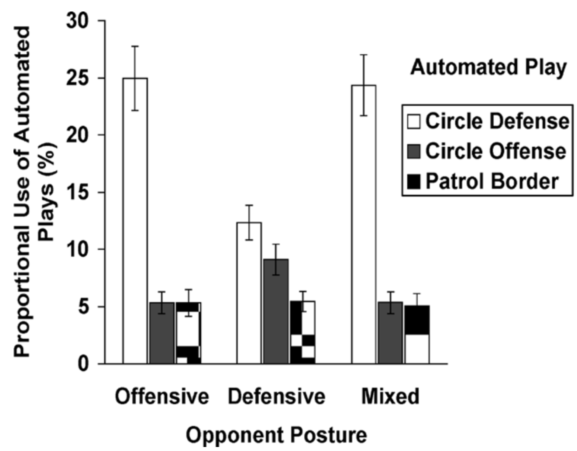


Fig. 4. Effects of opponent posture on proportional use of the three automated plays.

The longest missions were those where the participants played against the defensive stance ($M = 126.14$ s, $SE = 4.33$ s), which was also the condition when a 100% mission success rate was achieved. While participants had perfect success when going against a purely defensive opponent, they also took over twice as long to succeed. Conversely, participants had about a 75% success rate when they went up against an offensive opponent; however these missions were of the shortest duration. None of the other effects were significant.

It is natural to question whether the mission completion times were different depending on the mission outcome. The data shown in Fig. 2(c) indicate that the mission completion times were not significantly affected by mission success or failure for the offensive posture and only slightly different when the participants played against the mixed posture.

2) Strategy Usage: There were nine different states a robot could transition to/from: inactive, unassigned, circle defense, circle offense, patrol border, manual control, tagged, flagged circle offense, and flagged manual. The percent of time that the robots were under manual control was submitted to a 3 $\times$ 3 ANOVA. There was a significant interaction between Visual Range and Opponent Posture, $F(4, 68) = 2.96$, $p < 0.05$. As Fig. 3 shows, manual control was used less frequently against a defensive posture and more frequently under the offensive and mixed conditions. Furthermore, visual range mediated the use of manual control. When the visual range was low, indicating a higher uncertainty about opponent robot positions, manual control was used more often in the mixed condition. This was not the case in the offensive or defensive conditions.

We also analyzed the amount of time the robots were directed to perform a play. All plays were grouped so that the percent of time the robots were functioning in a particular state could be determined. This percent metric was submitted to a 3 $\times$ 3 ANOVA. There was a significant effect of Opponent Posture, $F(2, 34) = 11.34$, $p < 0.05$. Plays, regardless of type, were used most often and equally when the opponent stance was offensive or mixed, but less often with a defensive opponent. In a further analysis, the three plays (*circle defense, circle offense, patrol border*) were included as an additional factor in the statistical analysis, yielding a $3 \times 3 \times 3$ ANOVA. There was a significant interaction between the Opponent Posture and the type of play used, $F(4, 68) = 8.30$, $p < 0.05$. Fig. 4 illustrates that the *circle defense* play was the most utilized play compared to

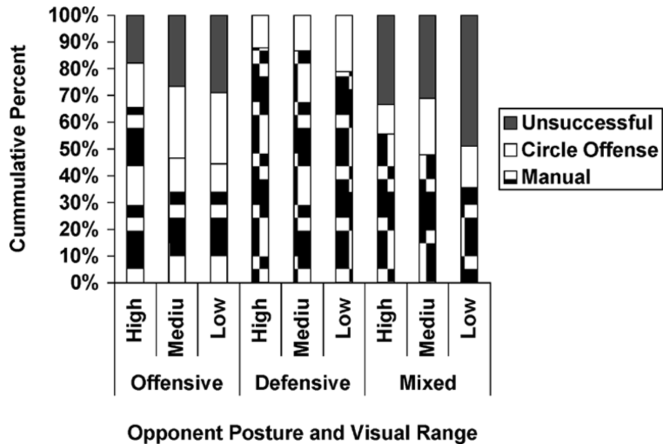


Fig. 5. Relative proportions of successful and unsuccessful missions, and the robot control state for successful missions (*circle offense* or manual).

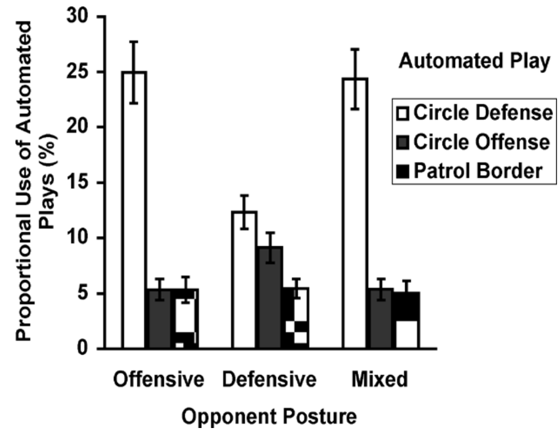


Fig. 6. Proportional use of manual control strategy to capture the flag on successful missions, as a function of opponent posture and robot visual range.

the other available plays, regardless of the opponent posture encountered. When the participants were faced with an offensive or mixed posture opponent, they used the *circle defense* play far more often than the *circle offense* or *patrol border* plays. While the proportion of automated plays was similar in the offensive and mixed conditions, the proportion was altered in the defensive condition. Participants utilized the *circle defense* play less often when they confronted a defensive opponent and increased the use of the *circle offense* play.

Another aspect of user strategy concerns the capturing of the opponent's flag. Participants could either use the *circle offense* play or direct robots manually to capture the flag and bring it to the midline to complete the mission successfully. Alternatively, the mission could end (unsuccessfully) with the opponent team crossing the midline with the participant's flag. Fig. 5 shows the three possible outcomes for ending a game. There was a potential nonindependence assumption violation in the factors so a statistical analysis could not be performed. Rather, Fig. 5 provides a profile of the percent of successful and unsuccessful missions, as well as the strategy that was employed for the robot that won each game under the nine possible experimental conditions. It is evident that participants preferred to use manual control rather than the *circle offense* play for the robot that captured the flag and returned midfield to end the mission.

The profile examines only the *end status* of each mission. A more sensitive measure that captures differences in strategies is the percent of time the robots were in the "Manual Flagged" state, which includes all instances in which the participant succeeded in the mission but also includes times when the participant captured the opponent's flag but was unable to cross the midfield line. The difference in the two measures is subtle but important. This measure illustrates the use of a strategy (in this case, manual control) regardless of the mission outcome. A 3 × 3 ANOVA revealed that there was a significant interaction between Visual Range and Opponent Posture for the percent of time participants chose manual control to capture the flag, $F(4,68) = 2.66$, $p < 0.05$. As Fig. 6 illustrates, participants utilized manual control most often when the opponent posture was defensive, followed by the mixed posture and finally the offensive posture. Regardless of the posture, participants used manual control to capture the opponent's flag most often when the visual range was high. Manual control usage under low and medium Visual Range was not remarkably different when the opponent posture was offensive or defensive but was clearly demarcated when the posture was mixed. This analysis shows that the specific act of acquiring the flag instead of traversing the field seems to have differentially affected the use of the manual control strategy.

3) Mental Workload and Situation Awareness: The NASA TLX and 3-D SART were administered after each block (visual range) of 15 trials. Overall subjective mental workload was computed by averaging the six NASA-TLX subscales and submitted to a three-level, one-way ANOVA. There was a significant difference in subjective mental workload, $F(2,34) = 8.77$, $p < 0.001$, across the low, medium, and high visual range conditions, with mental workload increasing as the visual range was reduced. There was no significant difference between the subjective assessments of overall situation awareness across the visual range conditions.

C. Discussion

Experiment 1 showed that the simplified delegation interface allowed for effective user supervision of robots, as evidenced by the number of missions successfully completed and time for mission execution. As expected, significantly fewer missions were successfully completed when the opponent posture was mixed rather than offensive. Nevertheless, users still had moderate success (about 62%) and relatively short mission completion times (about 51 s) in the mixed posture condition, which was the most challenging because of increased opponent uncertainty. These findings suggest, but do not prove, that delegation-type interfaces, as exemplified by the Playbook interface [9], [10], allowed users to respond effectively to unpredictable changes in opponent posture by tasking robots appropriately. Further confirmation of this view requires studies in which more complex versions of this type of interface are evaluated and compared against less flexible (nonadaptable) interfaces. Nevertheless, the findings add to the growing body of evidence pointing to the efficacy of adaptive/adaptable interfaces [28], [30].

Adaptive automation is thought to allow for regulation of operator mental workload and maintenance of situation awareness [28], [30], [43]–[45]. With respect to mental workload, there was an expected effect of visual range. As the robot vision radius was reduced, uncertainty about opponent robot positions and tactics increased and users reported greater overall workload. However, the experimental design we used did not allow us to examine the effects of opponent posture on workload. Presumably users found it more difficult to complete their mission when the opponent stance was mixed rather than purely offensive. It would be interesting to examine in a future study whether the Playbook interface would allow users to "balance out" the variations in their mental workload in response to changes in task demands in supervising multiple robots, as has been reported in studies with other adaptive interfaces [45].

Of the different user strategies, the most general finding was that manual control was used less frequently against a defensive opponent than against an offensive or mixed posture. In addition, when uncertainty was high due to low robot visual range, manual control was used more often in the mixed posture, but not the case in the offensive or defensive postures. The combination of low visual range and a mixed opponent posture represented the most challenging condition to the user. The greater use of manual control may reflect a greater perceived need for redirecting robots from previously assigned plays in this case, possibly due to the perception of the limitations of the automation to achieve the mission goal.

Experiment 1 provides a preliminary empirical evaluation of the use of delegation-type interfaces [9], [10] for human supervision of multiple robots, since we showed that subjects made appropriate use of many types of control methods depending on context. The results are suggestive of the effectiveness of this interface in allowing users flexible use of automation in response to changing task demands.

III. Experiment 2

Experiment 1 provided an initial empirical demonstration of the effectiveness of a delegation interface for supervision of multiple UVs. However, unlike previous research on adaptive or flexible automation [8], [28], a comparison to static delegation was not made. Accordingly, in Experiment 2 we compared the flexible delegation approach to *fixed* delegation approaches, by providing users: 1) *only* manual control or 2) *only* automated plays or 3) *both* types of control, under the same varying adversary postures (offensive, defensive, mixed) manipulated in Experiment 1.

We hypothesized that the use of the delegation interface would afford users maximum flexibility, allowing them to decide when workload was high (and therefore to use automation), or when the automation was ineffective. Additionally, we anticipated that the delegation interface would allow users to mount a more effective response to variable opponent postures than static control (manual or automated). As in Experiment 1, we tested these hypotheses by measuring overall mission performance indicators (success rate and time to mission completion) and operator mental workload and situational awareness.

A. Methods

1) Participants: Five males and four females between the ages of 19 and 33 ($M = 24.0$ years) participated.

2) Apparatus and Procedures: These were identical to those described in Experiment 1 with the exception that robot visual range was constant (medium) and that the interface was varied to be: 1) fixed, manual only; 2) fixed, automated plays only; or 3) flexible, both manual/automated plays. In the manual only condition, play selection was not available to the operator, who had to rely solely on waypoint control. In the automated plays condition, the operator could select any one of three plays available (*circle offense, circle defense, patrol border*) but was unable to use manual control (i.e., waypoint commands). When both options were available, the operator could use either method flexibly, at will. The opponent posture was also varied as in Experiment 1 to be offensive, defensive, or mixed.

A within-subjects design was employed, with three Delegation Types (fixed manual, fixed automated, flexible, manual/automated) combined with Opponent Posture (offensive, defensive, mixed), yielding nine conditions. Each participant completed five mission trials for each condition for a total of 45 trials. Delegation Type was a blocked factor while Opponent Posture was randomized within each block.

B. Results

1) Overall Performance: There was a significant effect of Opponent Posture on mission success rate, $F(2, 16) = 17.51$, $p < 0.01$. Expectedly, the participants won 100% of the games when the opponent strategy was defensive and the red team did not make a move to capture the blue team flag. There was no significant difference ($p > 0.05$) between the offensive and mixed conditions, where participants had success rates of 78% and 79%, respectively. There were no other significant differences ($p > 0.05$) for this measure.

Opponent Posture significantly affected mission completion times, $F(2, 16) = 56.61, p < 0.01$. Participants confronting an offensive red team completed the mission faster ($M = 31.87$ s, $SE = 0.51$ s) than against a mixed ($M = 39.27$ s, $SE = 1.79$ s) or defensive posture ($M = 103.47$ s, $SE = 6.44$ s). Further, there was a significant main effect of Delegation Type on mission completion times, $F(2, 16) = 4.88, p < 0.05$. The longest times were in the fixed automation condition ($M = 69.74$ s, $SE = 4.72$ s) compared to the fixed manual only ($M = 54.06$ s, $SE = 5.25$ s) and shortest times in the flexible manual/automated condition ($M = 50.81$ s, $SE = 4.09$ s), though these latter two did not differ statistically from one another.

2) Strategy Usage: The percentage of time the robots were commanded to use a particular play was included in a 3 (Opponent Posture) X 3 (Play used) X 2 (Delegation Type) ANOVA. There was a significant interaction between these factors, $F(4, 32) = 17.28$, $p < 0.01$. As expected, more automated plays were used when plays were the only method of robot supervision (see Fig. 7). However, given a choice, participants did not utilize automated plays to the degree that they could have. Further, the pattern of usage was consistent; *circle defense* was used the most often followed by *circle offense* and *patrol border*. The only exception to this pattern was in the automated only condition when participants faced a purely defensive red team, in which case they relied more on the use of the *circle offense* play.

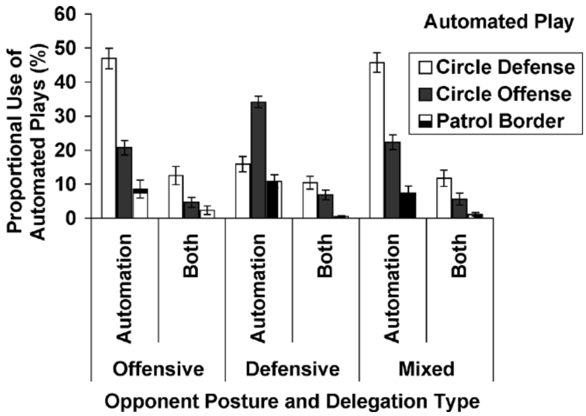


Fig. 7. Effects of delegation type (fixed automation or flexible, both manual/automated) and opponent posture on proportional use of automated plays.

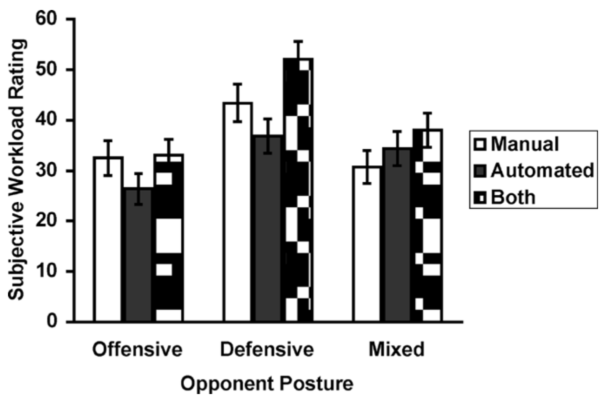


Fig. 8. Effects of delegation type and opponent posture on subjective mental workload.

Conversely, the percentage of time that robots were under manual control in the manual only condition compared to when both types of control were available was submitted to a 3 (Opponent Posture) X 2 (Delegation Type) ANOVA. There were significant main effects for Opponent Posture, $F(2,16) = 38.77$, $p < 0.01$, and Delegation Type, $F(1,8) = 8.26$, $p < 0.05$. Operators used manual control most often when playing against the red team offensive posture (67.23%) followed by the mixed condition (66.1%) and used manual control least often when playing the defensive red team strategy (53.7%). Also, participants used manual control 71.7% of the time when only manual control was available compared to 53.0% of the time when both automation and manual control were available.

3) Mental Workload and Situation Awareness: For the subjective rating of mental workload, there was a significant interaction between Opponent Posture and Delegation Type, $F(4,32) = 3.92$, $p < 0.05$. This interaction, illustrated in Fig. 8, indicates that participants rated their mental workload highest when they had both types of control available. Fig. 8 also shows that mental workload was rated higher in manual control versus automated control in all except the mixed red team posture, where the trend was reversed. Note, however, that the increase in mental workload in the flexible condition was relatively small, and the absolute values were in the low to moderate range (30–55, in a scale from 0–100).

For the situation awareness rating, there was a significant main effect for each of the factors—Opponent Posture, $F(2,16) = 5.49$, $p < 0.05$, and Delegation Type, $F(2,16) = 7.02$, $p < 0.01$. Participants rated their situation awareness highest when the red team status was offensive ($M = 78.30$, $SE = 1.65$) followed by the mixed posture ($M = 73.93$, $SE = 1.84$) and reported the lowest rating when playing against the defensive stance ($M = 71.19$, $SE = 1.88$). Furthermore, participants reported a higher level of situation awareness for the manual ($M = 82.00$, $SE = 1.27$) than the automation only condition ($M = 70.63$, $SE = 2.1$) or flexible conditions ($M = 70.78$, $SE = 1.78$).

C. Discussion

Operator usage of the delegation interface when flexibility was allowed (the flexible manual/automation condition) was different than the manual only or automation only conditions, as revealed by strategy usage of manual and automated control. Further, participants used the delegation interface to adapt to unpredictable opponent postures, as shown by the consistent defensive strategy to oppose forces when they were in an offensive or mixed posture, and alternated offensive strategy usage when no opposing robots were sent, as in the defensive posture. Even with the limited delegation interface used in this study, participants clearly were able to adapt effectively to the situation, as shown by the high rate of mission success (>75%). Manual control use in the flexible condition allowed participants the ability to overcome less effective automation behavior, thereby decreasing mission completion time, making it similar to the manual only condition and shorter than the automation only condition.

Another proposed benefit to delegation interfaces is the ability to off-load tasks when mental workload is increasing, or to intervene in robot actions if automated behavior is suboptimal [9], [10]. Expectedly, situation awareness was highest in the manual only condition as a result of decreased unpredictability. Increased opponent posture difficulty (indicated by mission completion time) resulted in lower situation awareness for the defensive posture. Interestingly, in the flexible manual/automation condition, participants did not retain the situation awareness benefits of increased robotic interaction, as previously described for mission completion time. This could be due to the small increase in mental workload associated with using the flexibility to decide between when to use automation or manual control.

Note that the tradeoff between situation awareness and workload is complex. While the delegation approach [9], [10] predicts increased situation awareness about what automation is doing (over fully automated approaches), the coarse-grained, subjective situation awareness measure used in these experiments was not focused on this parameter, but instead measured general situation awareness. It is therefore quite possible that the added workload required to decide and shift between control strategies resulted in less capacity to maintain general situation awareness relative to the automation only condition.

These results suggest an important conclusion. Automation brittleness or adversary tactics may require manual control, but this option involves a degree of workload, which could limit its use. The Playbook interface allows an operator to choose an operating point in this continuum of a tradeoff between the need for intervention and increased workload. The operator has flexibility in determining when automation is ineffective and can switch strategies when needed. By explicitly comparing a flexible delegation interface against static control, as in previous studies of adaptive automation [28], [32], the results of Experiment 2 also confirm that the increased flexibility of user-adaptable interfaces such as Playbook results in system and human performance benefits.

IV. EXPERIMENT 3

Experiment 2 showed that when operators had flexible access to *both* automated plays and manual control, mission completion times were reduced compared to a less flexible interface in which operators could only use automation. However, there was an associated cost with this flexibility in the form of a small increase in subjective mental workload. The third experiment further investigated the dimensions of flexibility offered by the Playbook interface. Of particular interest were two high-level dimensions of adaptation flexibility: *abstraction* and *aggregation*. Abstraction can be thought of as variation along a dimension of a task hierarchy, where primitive robotic behaviors, such as waypoint-to-waypoint movement, are at the lower level of abstraction. More complex behaviors such as *patrol border* or *circle defense* (continuous planned movement action and reaction to events such as opponent attack) are at a higher level of abstraction. An even higher level of abstraction would be to task a group of robots to go on *offense*—which implies some mix of offensive behaviors or plays. Aggregation can be defined as the number of robotic agents to which particular tasks are assigned as a group. Low aggregation refers to commands given to individual agents, whereas high aggregation means tasking all agents with the same plays. An intermediate, and flexible, level of aggregation is also possible where tasks can be given to groups of robots smaller or equal to that of the whole team.

In Experiment 3 we varied the number of robots supervised by the operator and had them compete against a similarly equipped opponent (four versus four or eight versus eight robots) using eight different interfaces representing different combinations of the abstraction and aggregation dimensions. Four of these conditions represented flexible (5–8), and the remaining four restricted interfaces (1–4), as defined and identified in Fig. 9: restricted interfaces are represented by the matrix cells within the boxed lines, flexible interfaces are outside the box, in the last column and bottom row. A full factorial combination of these factors was not used, rather, only those conditions that were important for our preplanned comparisons were included (numbered cells in Fig. 9). We hypothesized that an increased number of robots to supervise would reduce mission success and increase workload. It was expected that the flexibility of the RoboFlag interface would allow operators the ability to decide when workload was high (in which case, use automation), or when the automation was not effective due to limited observability or unpredictable opponent tactics (use manual control of robots, thereby decreasing unpredictability) and adjust accordingly. These benefits should not be available to operators using more restricted interfaces.

A. Methods

1) Participants: Three males and nine females between the ages of 18 and 27 ($M = 21.1$ years) participated.

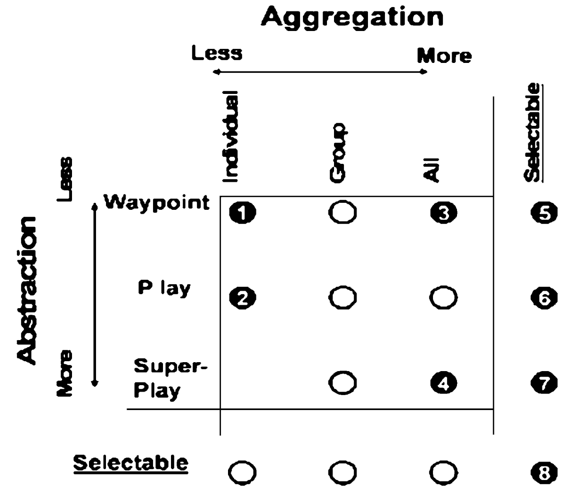


Fig. 9. All possible interface combinations of the dimensions of *abstraction* and *aggregation*. Interfaces 1–4 (restricted) and 5–8 (flexible) were chosen for investigation in Experiment 3.

2) Apparatus and Procedures: Of the eight interfaces selected, four represented restricted (1–4) and four flexible interfaces (5–8). The three levels of abstraction (y axis, Fig. 9) examined included: waypoint control, plays (*circle defense, circle offense, patrol border*), and superplay (higher level group plays comprised of more than a single play, generally allocating multiple robots across plays to provide an overall "posture," options being *offense, defense, mixed*). The three levels of aggregation (x axis, Fig. 9) available were the selection of individual (only one robot at a time), a subgroup (any number of robots, from one to all), or only all robots. The numbered cells in Fig. 9 show the selected interfaces that combined different levels of abstraction and aggregation. According to this taxonomy, Interface 8 had flexible control for both factors, whereas the other seven were progressively less flexible. Participants used these interfaces to supervise their robot team against an opponent robot team of equal force (either four or eight robots). The opponent posture was mixed, as defined previously. A within-subjects design was employed using eight interfaces selected from the possible options shown in Fig. 9. These were combined factorially with the number of robots controlled (four or eight). Interface type was treated as a blocked factor, with the number of robots randomized within each block. Each participant completed two sessions of four blocks each with ten trials for each block (five trials with four robots, five with eight robots) for a total of 80 trials.

B. Results

1) Number of Robots: The number of robots controlled was a significant factor in mission success rate as seen in Fig. 10, $F(1, 11) = 14.24$, $p < 0.01$. The ANOVA indicated that participants had more successful mission outcomes when they were controlling four robots ($M = 54.37\%$, $SE = 2.28\%$) than when they controlled eight robots ($M = 36.04\%$, $SE = 2.19\%$). Similarly, mission completion times were 8.27 s longer when eight robots were controlled compared to four robots, $F(1, 11) = 43.81, p < 0.01$.

2) Restricted versus Flexible Interfaces: There was a significant two-way interaction between the Number of Robots

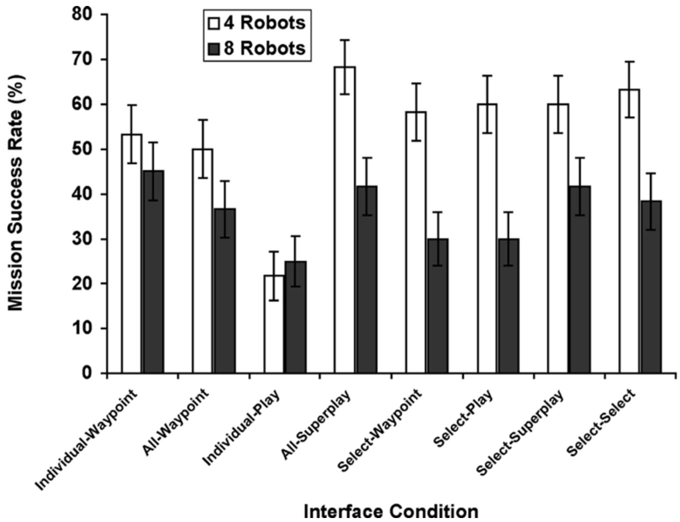


Fig. 10. Mission success rates for the eight interfaces representing combinations of the abstraction and aggregation dimensions.

controlled and the Interface Type for mission success rates, $F(1,11) = 11.56$, $p < 0.01$. This interaction, depicted in Fig. 11(a), shows that when there were four robots in the team, mission success rates were significantly higher with the flexible interfaces than with the restricted interfaces. However, this benefit for the flexible interfaces was not obtained when participants had to control eight robots. In fact, the success rate was virtually unchanged between the interface conditions when eight robots were controlled.

For mission completion times, there was a significant main effect for the number of robots, $F(1,11) = 43.81$, $p < 0.001$. Missions were completed an average of 8.26 s faster when participants were controlling only four robots compared to eight robots, virtually mirroring the effects found previously.

3) Restricted Interfaces: Analyses within the restricted interface conditions where the level of aggregation was "All" (interfaces 3 and 4) gave a significant interaction between the number of robots controlled and the two levels of aggregation for mission success rate, $F(1,11) = 20.79, p < 0.001$. This interaction, illustrated in Fig. 11(b), reveals that when participants had to assign all of the robots the same command, they achieved greater success when superplay was used compared to when only waypoint control was available. Further, while the number of robots controlled did not affect the outcome when waypoint control was available, the success rate was much higher when participants controlled four robots compared to eight robots when utilizing the superplay option. There was also a significant two-way interaction for this comparison for mission completion time, $F(1,11) = 10.79, p < 0.01$. As illustrated in Fig. 11(c), the missions that had only waypoint control available for the "All" condition robots were shorter than the missions that had only superplay available. Additionally, the number of robots controlled had opposite effects. Mission completion times were shorter in the waypoint condition when participants controlled eight robots compared to when they controlled four robots. Conversely, mission completion times were longer in the superplay condition when participants controlled eight robots compared to four. The percentage of successful missions

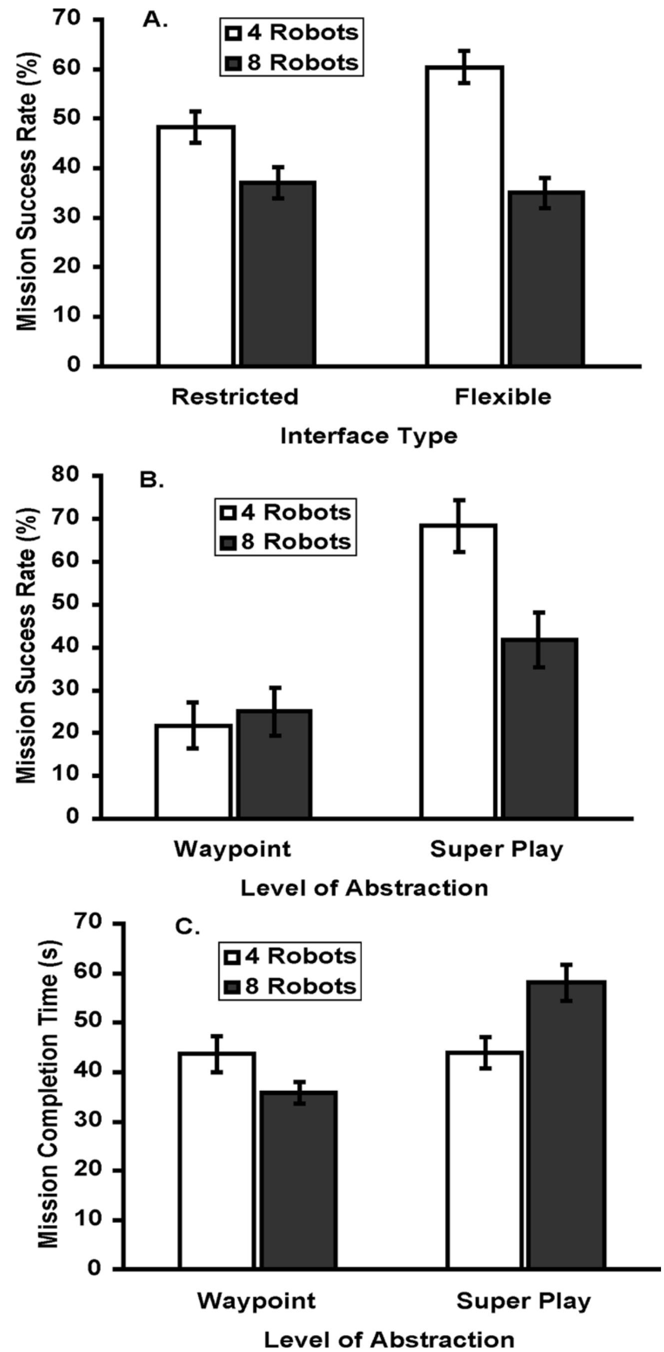


Fig. 11. Effects of number of robots and (a) interface type (restricted, flexible), and of level of abstraction (waypoint, super play) on (b) mission success rate and (c) mission completion time.

and the completion times suggest that participants were able to achieve a higher success rate using the superplay over manual control given that they were controlling all of the robots.

4) Flexible Interfaces: There were no significant differences between the different subtypes of flexible interfaces. There were differences for the mission success rate and mission completion times but only for the main effect of Number of Robots. The percentage of successful mission outcomes nearly doubled when the number of robots controlled was reduced from eight (35.0%) to four (60.4%) under the flexible interfaces, $F(1,11) = 17.50$, $p < 0.01$. Similar to the previous

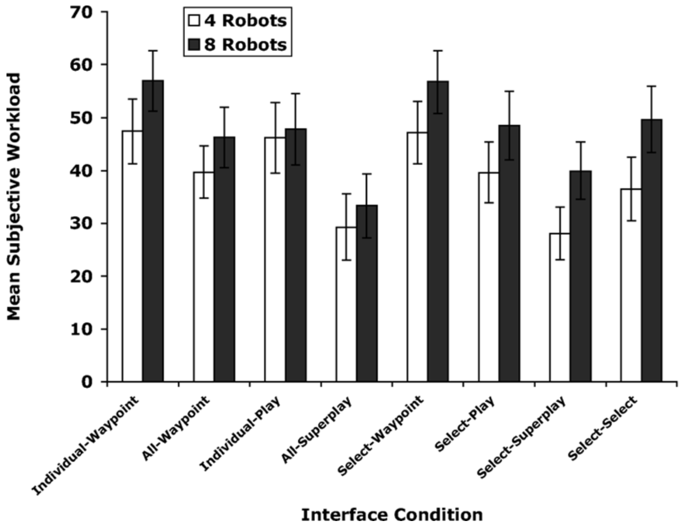


Fig. 12. Mean subjective workload for the eight interfaces representing combinations of the abstraction and aggregation dimensions.

results, the mission completion times were significantly shorter, $F(1,11) = 17.39$, $p < 0.01$, when four robots were controlled (42.19 s) compared to eight robots (52.37 s).

5) Mental Workload and Situation Awareness: Analysis of the subjective mental workload ratings revealed that the main effects of Number of Robots, $F(1,11) = 23.67$, $p < 0.001$, and Interface Type, $F(7,77) = 7.83$, $p < 0.001$, were significant, as was the interaction, $F(7,77) = 2.36$, $p < 0.05$. Planned comparisons revealed that participants reported significantly higher mental workload, $t(11) = 4.86, p < 0.001$, when supervising eight robots ($M = 47.39$, $SE = 5.38$) compared to four robots ($M = 39.23$, $SE = 5.15$) [see Fig. (12)]. Also, participants reported higher workload $t(11) = 5.28, p < 0.001$ for conditions where only waypoint control was available (individual waypoint, all waypoint, flexible waypoint) ($M = 50.38$, $SE = 5.73$) than when automated plays were available (individual play, all superplay, flexible play, flexible superplay) ($M = 38.07$, $SE = 5.26$).

For situation awareness, a significant main effect was obtained for Number of Robots, $F(1,11) = 21.87, p < 0.001$. Higher ratings were reported when four ($M = 74.71$, $SE = 3.37$) rather than eight ($M = 67.90$, $SE = 3.97$) robots were supervised. No other significant effects were found for this measure.

C. Discussion

Experiments 1 and 2 demonstrated that a delegation-type interface [9], [10] provides for effective human supervision of multiple autonomous robots, in comparison to nonadaptable interfaces. However, in the present study a more stringent test of the efficacy of flexible delegation was conducted by pitting it against a variety of more restricted interfaces. Overall performance was reduced when participants had to supervise eight as opposed to four robots. The results demonstrated that the effectiveness of Playbook-like interfaces stems primarily from the flexibility it affords the human user, compared to more restricted interfaces. We defined two dimensions pertaining to flexibility, level of aggregation and level of abstraction, and found that when operators controlled four robots the interface having the highest flexibility on these dimensions (the "Select–Select" interface) led to more successful missions and shorter mission completion time than the less flexible interfaces. This benefit was reduced, however, when eight robots were supervised. At this higher load, the performance benefit may have been countered due to the greater management workload demand imposed by full flexibility.

The findings also point to the potential problems associated with restricted interfaces (such as individual play, and all waypoint), which do not provide the operator the flexibility to reallocate robotic resources or to compensate for suboptimal robot behavior in response to unpredictable mission events. In general, the results provide increasing empirical evidence that the (most flexible) delegation-style interface provides benefits to a single operator, by allowing the operator the ability to adapt and respond to ineffective automation behavior at specific mission times during the control of multiple UVs.

V. Computational Modeling of Mental Workload

The use of flexible delegation interfaces was associated with efficient human–robot teaming performance, but at the cost of a slight increase in workload. We carried out a cognitive simulation analysis to account for the workload findings, using a task network model involving a dissection of RoboFlag into subtasks with sequential paths among the component tasks [46]. Following the task breakdown, a probabilistic model using a normal distribution estimated completion times for each subtask. In the next step, the cognitive resources associated with the completion of each subtask was simulated using the multiple-resource model of Wickens [47], in which five types of resources are defined: visual, auditory, cognitive, motor, and speech. Values for each cognitive resource were assigned in a model using a published database [48]. Monte Carlo simulation was then carried out to provide quantitative values for total momentary workload based on the estimated cognitive resources.

The task network model of the RoboFlag operator was constructed considering three fundamental steps of the operators' cognitive work: state recognition, decision-making, and operation. We considered two major factors in the model that would affect cognitive work across the different RoboFlag interface types. One was the probability of the need for the user to intervene manually in controlling robots in order to ensure mission success. We assumed that this probability would, by definition, be relatively high for the waypoint-only interface, lower for play operations, and much lower for superplay operations. Also, we assumed a higher probability when eight rather than four robots had to be supervised. The second factor was the operator's time for decision-making. We assumed that this would be shorter when the number of operational alternatives was small, and longer when the number was large. For example, in the Select–Select condition, an operator should choose from among three options, waypoint, play, or superplay, then select the number of robots to which the option should be applied, and finally execute the plan.

One thousand trials of Monte Carlo simulations were performed for each of the eight interface types examined in Experiment 3. To compare the results under the different condi-

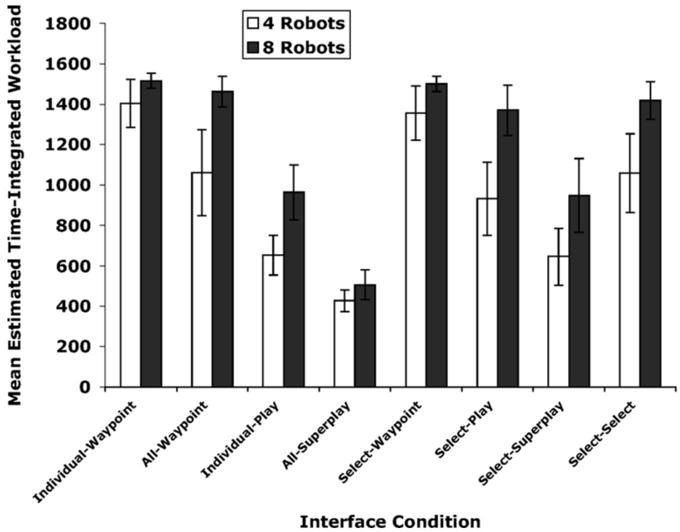


Fig. 13. Mean expected values of time-integrated workload from the simulation analysis for the eight interfaces.

tions, the simulated operational time was set equally at 60 s. Fig. 13 shows the mean expected values of the total time-integrated workload for each of the eight interfaces. As expected, workload was higher when eight rather than four robots were supervised. Also, workload was higher when only waypoint control was available (individual waypoint, all waypoint, flexible waypoint) compared to when automated plays could be used, with the exception that relatively low values were found for the "All–Waypoint" interface. Finally, workload was high in the "Select–Select" interface, particularly when eight robots were supervised. These simulation findings closely parallel those reported in the empirical data from Experiment 3 (see Fig. 12).

VI. Implementing Delegation-Type Interfaces for Unmanned Vehicle Control

The three experiments presented here provide the initial empirical support for the efficacy of delegation-type interfaces in supervising multiple UVs. Other ongoing work [49], [50] on the implementation and application of such interfaces provides further support for their use in complex human–robot systems. This work has been conducted in a high-fidelity simulation emulating multiple small UAVs (fixed and rotary wing) operating in an urban environment to perform useful support services for small units of infantry soldiers.

Sequential control of multiple air vehicles was emphasized, rather than concurrent and coordinated control as in the RoboFlag studies presented previously. For example, one "play" that has been implemented and tested is known as "Watch Area" and allows an operator to very quickly command that surveillance images be provided for a designated area within a given time interval. By making extensive use of moderately smart default values, this play can be coarsely commanded in as few as three actions (mouse clicks on a desktop PC or stylus taps on a PDA) and less than 15 s of interaction time. Of course, this leaves much of the planning to be done by the automation. The resulting plan, although guaranteed to adhere to the constraints imposed by the operator, might nonetheless not be exactly what the user wanted or would have chosen, in which case the user can change the plan. The Playbook interface also permits more detailed interactions and specification of constraints including the designation of specific start and end times for imagery, the area to be scanned and whether scanning must be continuous or may be intermittent. This developmental prototype is fully integrated with a high-fidelity flight control system and provides closed-loop control of our simulated vehicles, including the ability for the Playbook to monitor and adapt plans in response to disturbances which might disrupt travel times such as headwinds or engine malfunctions.

VII. Conclusion

When human operators are required to supervise multiple UVs, as in the present studies, individual operator control of all robots is difficult, mandating the use of automation. At the same time, limitations in the reasoning capabilities of semiautonomous agents and the brittleness of some automated behaviors necessitate intervention through human intelligence, indicating that the human–robot interface must support multiple or adjustable levels of interaction [26]–[30]. The results of the three experiments reported here provide the initial empirical evidence that delegation-type interfaces such as Playbook provide a flexibility that enhances the performance of the human–robot team in comparison to static or more restricted interfaces. While there is considerable empirical support for the performance benefits of adaptive automation [28], [30], [32], similar evidence for efficacy when human users task automation—adaptable automation—has been lacking until the present study. We propose that delegation type interfaces such as Playbook allow for such multi-level interaction in a flexible manner that keeps human workload within a manageable range. In turn, such interfaces can provide for the effective coordination between humans and robotic agents that is being sought in various operational contexts [51].

Delegation architectures allow the human supervisor to declare goals and provide instructions and thereby choose how much or how little autonomy to delegate to automation on an instance-by-instance basis. Delegation interfaces have additional features that provide for flexibility of supervision of automated agents [9], [10]. Only the first of these features, the ability to task robots at different levels of abstraction or depth, was examined in the current series of experiments. Future studies should examine these other features, such as the ability to provide goals or plans to the robots, or the possibility of constraining robot performance at different and/or heterogeneous levels of depth.

Acknowledgment

The authors thank M. Campbell for the RoboFlag software and J. Adams, M. Barnes, and J. Bay for useful comments and advice.

References

[1] National Research Council, *Technology Development for Army Unmanned Ground Vehicles*. Washington, DC: Nat. Acad. Press, 2002.

[2] Z. Zigoris, J. Siu, O. Wang, and A. Hayes, "Balancing automated behavior and human control in multi-agent systems: a case study in RoboFlag," in *Proc. Amer. Controls Conf.*, Denver, CO, 2003, pp. 667–671.

[3] J. A. Adams, "Critical considerations for human–robot interface development," in *Proc. Human–Robot Interaction, 2002 AAAI Fall Symp.*, Menlo Park, CA, 2002, pp. 1–8.

[4] R. Parasuraman, S. Galster, and C. Miller, "Human control of multiple robots in the RoboFlag simulation Environment," in *Proc. 33rd Annu. IEEE Conf. Systems, Man, Cybernetics*, Washington, DC, 2003, pp. 3232–3237.

[5] E. Rogers and R. Murphy, "Human–robot interaction: Final report for DARPA/NSF study on human–robot interaction," Comput. Sci. Dept., California Polytech. State Univ., San Luis Obispo, 2001. [Online]. Available: http://www.csc.calpoly.edu/~erogers/HRI/HRI-report-final.html.

[6] S. Kiesler and S. P. Hinds, "Special issue on human–robot interaction," *Human–Comput. Interact.*, vol. 19, no. 1–2, 2004.

[7] J. Scholz, "Theory and evaluation of human robot interaction," in *Proc. Conf. Human Interaction with Complex Systems*, Urbana-Champaign, IL, 2000, pp. 232–237.

[8] R. Parasuraman, "Designing automation for human use: Empirical studies and quantitative models," *Ergonomics*, vol. 43, pp. 931–951, 2000.

[9] C. Miller and R. Parasuraman, "Beyond levels of automation: An architecture for more flexible human automation collaboration," in *Proc. Human Factors and Ergonomics Soc., 47th Annu. Meeting*, Denver, CO, 2003, pp. 182–186.

[10] C. Miller, M. Pelican, and R. Goldman, "'Tasking' interfaces for flexible interaction with automation: Keeping the operator in control," in *Proc. Conf. Human Interaction with Complex Systems*, Urbana-Champaign, IL, 2000, pp. 188–202.

[11] D. J. Bruemmer, D. D. Dudenhoeffer, and J. Marble, "Mixed-initiative remote characterization using a distributed team of small robots," in *Proc. AAAI Mobile Robot Workshop*, Seattle, WA, 2001, pp. 101–106.

[12] H. Mirmohammad-Sadeghi, H. Bastani, and F. Azarnasab, "IUT Microbot Robocup," presented at the *2003 Int. Robocup Conf.*, Padova, Italy.

[13] P. Ryan, "Mine countermeasures a success," *Naval Inst. Proc. Mag.*, vol. 129, no. 5, May 2003. [Online]. Available: http://www.usni.org/Proceedings/Articles03/PROryan05.htm.

[14] M. Campbell, R. D'Andrea, D. Schneider, A. Chaudhry, S. Waydo, J. Sullivan, J. Veverka, and A. Klochko, "RoboFlag games using systems based hierarchical control," in *Proc. Amer. Controls Conf.*, Denver, CO, 2003, pp. 661–666.

[15] Army Science Board, *Ad Hoc Study on Human Robot Interface Issues*. Arlington, VA: Washington Headquarters Services, Directorate for Inf. Oper. Rep., 2002.

[16] R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," *Human Factors*, vol. 39, pp. 230–253, 1997.

[17] N. B. Sarter, D. Woods, and C. Billings, "Automation surprises," in *Handbook of Human Factors and Ergonomics*, G. Salvendy, Ed. New York: Wiley, 1997, pp. 1926–1943.

[18] R. Parasuraman and E. A. Byrne, "Automation and human performance in aviation," in *Principles of Aviation Psychology*, P. Tsang and M. Vidulich, Eds. Mahwah, NJ: Erlbaum, 2003, pp. 311–356.

[19] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 30, pp. 286–297, 2000.

[20] R. Parasuraman and C. Miller, "Trust and etiquette in high-criticality automated systems," *Commun. ACM*, vol. 47, no. 4, pp. 51–55, Apr. 2004.

[21] W. M. Crocoll and B. G. Coury, "Status or recommendation: selecting the type of information for decision aiding," in *Proc. Human Factors Soc.*, Santa Monica, CA, 1990, pp. 1524–1528.

[22] E. Rovira, K. McGarry, and R. Parasuraman, "Effects of unreliable automation on decision making in command and control," in *Proc. Annu. Meeting of the Human Factors and Ergonomics Soc., 46th Annu. Meeting*, Baltimore, MD, 2002, pp. 428–432.

[23] N. Sarter and B. K. Schroeder, "Supporting decision-making and action selection under time pressure and uncertainty: The case of in-flight icing," *Human Factors*, vol. 43, pp. 573–583, 2001.

[24] C. D. Wickens, "Imperfect and unreliable automation and its implications for attention allocation, information access and situation awareness," Univ. Illinois, Urbana-Champaign, Tech. Rep. ARL-00-10/NASA-00-2, 2000.

[25] M. Wilson and M. Neal, "Diminishing returns of engineering effort in telerobotic systems," *IEEE Trans. Syst., Man, Cybern., Syst. Humans*, vol. 31, pp. 459–465, 2001.

[26] J. Crandall and M. A. Goodrich, "Principles of adjustable interactions," in *Human–Robot Interaction, Proc. 2002 AAAI Fall Symp.*. Menlo Park, CA, 2002, pp. 29–38.

[27] N. Moray, T. Inagaki, and M. Itoh, "Adaptive automation, trust, and self-confidence in fault management of time-critical tasks," *J. Exper. Psychol., Applied*, vol. 6, pp. 44–58, 2000.

[28] R. Parasuraman, "Effects of adaptive function allocation on human performance," in *Human Factors and Advanced Aviation Technologies*, D. J. Garland and J. A. Wise, Eds. Daytona Beach, FL: ERAU, 1993, pp. 149–157.

[29] W. B. Rouse, "Adaptive aiding for human/computer control," *Human Factors*, vol. 30, pp. 431–443, 1988.

[30] M. Scerbo, "Adaptive automation," in *International Encyclopedia of Ergonomics and Human Factors*, W. Karwowski, Ed. New York: Taylor & Francis, 2001.

[31] R. Opperman, *Adaptive User Support*. Hillsdale, NJ: Erlbaum, 1994.

[32] R. Parasuraman, M. Mouloua, and R. Molloy, "Effects of adaptive task allocation on monitoring of automated systems," *Human Factors*, vol. 38, pp. 665–679, 1996.

[33] C. E. Billings and D. D. Woods, "Concerns about adaptive automation in aviation systems," in *Human Performance in Automated Systems: Current Research and Trends*, R. Parasuraman and M. Mouloua, Eds. Hillsdale, NJ: Erlbaum, 1994, pp. 264–269.

[34] T. Sheridan, "Supervisory control," in *Handbook of Human Factors*, G. Salvendy, Ed. New York: Wiley, 1987, pp. 1244–1268.

[35] K. Erol, J. Hendler, and D. Nau, "UMCP: A sound and complete procedure for hierarchical task network planning," in *AI Planning Systems: Proc. 2nd Int. Conf.*, K. Hammond, Ed., Los Altos, CA, 1994, pp. 249–254.

[36] J. Rasmussen, *Information Processing and Human–Machine Interaction*. Amsterdam, The Netherlands: North-Holland, 1986.

[37] S. R. Dixon and C. D. Wickens, "Control of multiple UAVs: a workload analysis," in *Proc. 12th Int. l Symp. Aviation Psychology*, Dayton, OH, Ap. 2003.

[38] H. A. Ruff, S. Narayanan, and S. M. H. Draper, "Human interaction with levels of automation and decision aid fidelity in the supervisory control of multiple simulated air vehicles," *Presence: Teleoper. Virtual Environ.*, vol. 11, no. 4, pp. 335–351, 2003.

[39] J. Ververka and M. Campbell, "Experimental study of information load on operators in semi-autonomous systems," in *Proc. AIAA Guidance, Navigation and Control Conf.*, Austin, TX, 2003, pp. 1–13.

[40] R. D'Andrea and M. Babish, "The RoboFlag testbed," in *Proc. Amer. Controls Conf.*, Denver, CO, 2003, pp. 650–655.

[41] S. G. Hart and LE. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds. Amsterdam, The Netherlands: Elsevier Science/North Holland, 1988, pp. 139–183.

[42] R. M. Taylor, "Situational awareness rating technique (SART): The development of a tool for aircrew systems design," in *Situational Awareness in Aerospace Operations*. Neuilly Sur Seine, France: NATO-AGARD, 1990. AGARD-CP-478.

[43] D. B. Kaber and M. R. Endsley, "The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task," *Theor. Issues Ergonom. Sci.*, vol. 5, no. 2, pp. 113–153, 2004.

[44] P. A. Hancock, M. H. Chignell, and A. Lowenthal, "An adaptive human–machine system," in *Proc. IEEE Conf. Systems, Man, Cybernetics*, 1985, pp. 627–629.

[45] B. Hilburn, P. G. Jorna, E. A. Byrne, and R. Parasuraman, "The effect of adaptive air traffic control (ATC) decision aiding on controller mental workload," in *Human–Automation Interaction*, M. Mouloua and J. Koonce, Eds. Mahwah, NJ: Erlbaum, 1997, pp. 84–91.

[46] K. R. Laughery, "Modeling human performance during system design," in *Human/Technology Interaction in Complex Systems*, E. Salas, Ed. Stamford, CT: JAI, 1999, pp. 147–174.

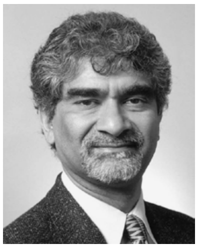
[47] C. D. Wickens and Y. Y. Yeh, "A multiple resource model of workload prediction and assessment," in *Proc. IEEE Conf. Systems, Man, Cybernetics*, 1986, pp. 1044–1048.

[48] Micro Analysis and Design, Inc., *Micro Analysis and Design, User's Manual of WinCrew: Windows-Based Workload and Task Analysis Tool*. Boulder, CO: Micro Analysis and Design, Inc., 1997.

[49] C. Miller, H. Funk, R. Goldman, and P. Wu, "A 'Playbook' for variable autonomy control of multiple, heterogeneous unmanned air vehicles," in *Human Performance, Situation Awareness, and Automation*, D. A. Vicenzi, M. Mouloua, and P. A. Hancock, Eds. Mahwah, NJ: Erlbaum, 2004, pp. 210–214.

[50] R. Goldman, C. Miller, P. Wu, H. Funk, and J. Meisner, "Optimizing to satisfice: using optimization to guide users," presented at the *Amer. Helicopter Soc. Int. Specialists Meeting on Unmanned Aerial Vehicles*, Chandler, AZ, 2005.

[51] D. D. Woods, J. Tittle, M. Feil, and A. Roesler, "Envisioning human–robot coordination in future operations," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 34, no. 2, pp. 210–218, May 2004.



Raja Parasuraman received the B.Sc. (first class honors) degree in electrical engineering from Imperial College, University of London, London, U.K., the M.Sc. degree in applied psychology and the Ph.D. dergee in psychology, both from Aston University, Birmingham, U.K., in 1972, 1973, and 1976, respectively.

He is currently a Professor of psychology at George Mason University, Fairfax, VA. Previously, he was Professor of psychology at The Catholic University of America. He has conducted research on human performance in complex human–machine systems, particularly with respect to the influence of automation on attention and decision-making. He also has an extensive program of research in cognitive neuroscience and in neuroergonomics. His research is these areas has been supported by several federal agencies, including NASA, NIH, DOD, and DARPA, as well as by private foundations. His books include *Automation and Human Performance* (Mahwah, NJ: Erlbaum, 1996), *The Attentive Brain* (Cambridge, MA: MIT Press, 1998), and *Neuroergonomics: The Brain at Work* (New York: Oxford Univ. Press, 2005). He is on the editorial board of *Human Factors* and *Theoretical Issues in Ergonomics Science*.

Dr. Parasuraman received the Franklin V. Taylor Award for Lifetime Achievement in Applied Experimental and Engineering Psychology from the American Psychological Association (Division 21) in 2004. He was also the recipient in 1997 and again in 2001 of the Jerome H. Ely Award for best paper in the journal *Human Factors* by the Human Factors and Ergonomics Society. He is Chair of the National Research Council Panel on Human Factors. He was elected a Fellow of the American Association for the Advancement of Science (1994), the American Psychological Association (1991), the American Psychological Society (1991), the Human Factors and Ergonomics Society (1994), and a National Associate of the National Academy of Sciences (2001).

Scott Galster received the B.A. degree in psychology from The Ohio State University, Columbus, and the Ph.D. degree in experimental psychology from The Catholic University of America, Washington, DC, in 1985 and 2003, respectively.

He is currently a Program Manager for the advanced research project for Multi-Sensor Command and Control Aircraft at the Air Force Research Laboratory, Wright-Patterson Air Force Base, OH. His research interests include human interaction with automation, battle management command and control systems, and advanced technology evaluations.

Peter Squire received the B.S. degree in computer science from Mary Washington College, Fredericksburg, VA, in 2002. He is currently working toward the Ph.D. degree in human factors and applied cognition at George Mason University, Fairfax, VA.

He is currently a Scientist for the B 40 Human System Integration Branch, Naval Surface Warfare Center Dahlgren Division (NSWCDD), Dahlgren, VA. His research interests are in adaptive automation, human–robot interaction, and neuroergonomics.

Hiroshi Furukawa received the B.E., M.E., and Ph.D. degrees in nuclear engineering from Tohoku University, Sendai, Japan, in 1990, 1992, and 1995, respectively.

He is currently an Associate Professor in the Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba, Japan. He was a Postdoctoral Researcher at the Human Factors Research Laboratory, Japan Atomic Energy Research Institute, from 1996 to 1998. His research interests include human interface designs for human–machine collaboration, quantitative and qualitative evaluation methods for function allocation among humans and automated systems, and formation process of navigational knowledge.

Christopher Miller received the B.A. degree from Pomona College, Claremont, CA, and the M.A. and Ph.D. degrees from the University of Chicago, Chicago, IL, in 1985 and 1991, respectively, all in psychology.

He worked at the Honeywell Technology center from 1989 to 2001 in various positions culminating in Research Fellow. He is currently Chief Scientist and co-owner of Smart Information Flow Technologies, a small business emphasizing research and development in adaptive automation and information management. He is also currently serving as an Adjunct Professor in the Human Factors program at the Department of Kinesiology, University of Minnesota, Duluth. He is the author of over 120 journal articles, book chapters, and conference proceeding papers on adaptive automation and related topics. He currently serves on the editorial boards of *Human Factors* and *Knowledge-Based Systems*.