NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer’s or trade names does not constitute an official endorsement or approval of the use thereof.

DESTRUCTION NOTICE—Destroy this report when it is no longer needed. Do not return it to the originator.
Implementation of Automation for Control of Robotic Systems

Keryl A. Cosenzo and Michael Barnes
Human Research and Engineering Directorate, ARL

Raja Parasuraman and Anthony Novak
George Mason University

Approved for public release; distribution is unlimited.
Future Combat Systems (FCS) represent an integrated materiel approach to transform the future U.S. Army into a more lethal, deployable, and survivable force. Technology alone will not ensure efficacy. Soldiers (and specifically their performance in this environment) will determine the success or failure of the fielded systems. In particular, robotic technology will be a vital component of future combat because it will extend manned capabilities, acting as force multipliers, and most importantly, it will save lives. The role of the human operator in the human-robot environment is not well understood; however, most contemplated systems will require active human control or supervision with the possibility of intervention. In the most extreme case, Soldiers will operate multiple systems while moving and while undergoing enemy fire. In all cases, workload and stress will be variable and unpredictable, changing rapidly as a function of the military environment.

Automation technologies have been successfully applied to aid human operators in various environments, including aviation and military command and control. This report addresses strategies to minimize the demands on Soldiers in the robotic environment through the use of adaptive and adaptable automation. Adaptable interfaces allow the Soldier to define conditions for automation decisions during mission planning while adaptive interfaces automate tasks as a function of some environmental or behavioral indicator (Parasuraman, Sheridan, & Wickens, 2000). Although multiple robot control and the application of adaptive and adaptable automation have been investigated in some contexts, they have not been investigated as an aid to multiple robot control. We are examining the use of adaptive or adaptable automation to assist an operator who will control multiple robotic aerial and ground systems from a single interface in a vehicular environment. In this report, we provide an overview of the current state of the research. We also discuss the joint research being conducted by the U.S. Army Research Laboratory and George Mason University on adaptive and adaptable interfaces that unload the Soldier during overload or emergency situations for the robotic multitasking environment. The results of the first experiment are reported as a baseline assessment of task performance in a simulated robotic environment without automation. We discuss the implications of these results for designing adaptive systems and the future directions for this research.
## Contents

### List of Figures

<table>
<thead>
<tr>
<th>1. Introduction</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Robotic Systems in the Military</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Automation</td>
<td>3</td>
</tr>
<tr>
<td>1.2.1 Levels of Automation</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Human Performance Issues for Automated Systems</td>
<td>4</td>
</tr>
<tr>
<td>1.3.1 Trust in Automation</td>
<td>4</td>
</tr>
<tr>
<td>1.4 Additive Principles</td>
<td>7</td>
</tr>
<tr>
<td>1.4.1 Characteristics of Adaptive Automation Systems</td>
<td>8</td>
</tr>
<tr>
<td>1.4.2 Adaptive Versus Adaptable Automation</td>
<td>10</td>
</tr>
<tr>
<td>1.4.3 Human Interaction With Adaptive Systems</td>
<td>11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Research Approach</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Robotic NCO Program</td>
<td>13</td>
</tr>
<tr>
<td>2.2 Program Specifications</td>
<td>13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Experiment I</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Participants</td>
<td>14</td>
</tr>
<tr>
<td>3.2 Apparatus</td>
<td>14</td>
</tr>
<tr>
<td>3.2.1 Robotic NCO</td>
<td>14</td>
</tr>
<tr>
<td>3.2.2 Workload and Situation Awareness Questionnaires</td>
<td>15</td>
</tr>
<tr>
<td>3.3 Experimental Scenario</td>
<td>15</td>
</tr>
<tr>
<td>3.4 Procedure</td>
<td>16</td>
</tr>
<tr>
<td>3.5 Experimental Design</td>
<td>17</td>
</tr>
<tr>
<td>3.5.1 UGV</td>
<td>17</td>
</tr>
<tr>
<td>3.5.2 UAV</td>
<td>17</td>
</tr>
<tr>
<td>3.5.3 Communications</td>
<td>18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Results</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 UGV Task Performance—Reaction Time (RT) for AOIs and Obstacles</td>
<td>19</td>
</tr>
<tr>
<td>4.2 UAV Task Performance—RT to Targets and Percent Correctly Identified</td>
<td>20</td>
</tr>
<tr>
<td>4.3 COMM Task Performance—Percent of Missed Call Signs and RT to Communications</td>
<td>20</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>4.4</td>
<td>Self-Reported SA Ratings</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Overall Situation Awareness</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Levels of Situation Awareness</td>
</tr>
<tr>
<td>4.5</td>
<td>Overall Workload</td>
</tr>
<tr>
<td>5.</td>
<td>Conclusions</td>
</tr>
<tr>
<td>5.1</td>
<td>Findings from Experiment I and Future Experimentation</td>
</tr>
<tr>
<td>5.2</td>
<td>Candidate Physiological Measures</td>
</tr>
<tr>
<td>5.3</td>
<td>General Conclusions</td>
</tr>
<tr>
<td>6.</td>
<td>References</td>
</tr>
<tr>
<td></td>
<td>Distribution List</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1. Hypothetical information flow. ................................................................. 2
Figure 2. Example of a closed loop adaptation for A, automated; A/M, automated/manual; and M, manual task sets. ............................................................................................................. 7
Figure 3. Prototype operator control unit............................................................... 13
Figure 4. Robotic NCO script creator................................................................. 14
Figure 5. Robotic NCO screen shots................................................................. 16
Figure 6. Mean reaction times (seconds) for obstacles in the UAV and UGV conditions. .... 19
Figure 7. Mean RT (seconds) for communication events in the low COMM condition. .... 20
Figure 8. Mean reaction time (seconds) for communication events in the high COMM condition. ................................................................................................................................. 20
Figure 9. Mean situation awareness in the low UAV condition ....................... 22
Figure 10. Mean situation awareness in the high UAV condition ....................... 22
Figure 11. Mean correct for type 1 SA communications in the UAV and COMMS conditions ................................................................................................................................. 23
1. Introduction

1.1 Robotic Systems in the Military

Future Combat Systems (FCS) is a U.S. Army program that will transform the battlefield. Future force structure, doctrine, and tactics will change as new systems are introduced, possibly in ways that cannot be anticipated. Elements of the force structure are being designed to be flexible, reconfigurable components of FCS tailored to specific combat missions. One aspect of increased flexibility will be the introduction of numerous robotic systems. The term robot is used in a generic sense to describe systems that are unmanned with some degree of autonomy and includes aerial, ground, subterranean, naval surface, and subsurface vehicles. These systems will be an essential part of the future force. It is anticipated that they will extend manned capabilities, be force multipliers, and (most important) save lives.

Robotic assets will be used as part of a team. For example, the robotic asset may be sent ahead of a troop to provide information about location and size of enemy units so as to facilitate safe movement of the troop. Robots may be given firing capabilities so that the Soldier is able to engage an enemy without being present, thereby increasing survivability. Currently, robots are being used to identify and remove unexploded ordnance. Larger robotic assets such as the unmanned air vehicle (UAV) can provide reconnaissance information about areas far away from a unit. Even though the addition of assets may enhance the lethality and survivability of the U.S. military, this benefit does have a cost. For example, the current UAVs require multiple operators to control them, and these operators have no other tasks to complete. Smaller ground systems are currently teleoperated by an operator and again, this is the operator’s primary task. In the future, it is anticipated that robotic control will be one of many tasks the Soldier will be responsible for. The future Soldier will conduct “traditional military tasks” such as scanning for threats, engaging targets, and conducting communications. Additionally, the Soldier will have to control or monitor the unmanned system(s) and process information returned from the unmanned system(s).

It is possible that the robotic control task and the information processing task for the unmanned system(s) are done by different individuals, which will minimize some of the workload. The robotic operator typically will not be the decision maker; s/he will send a situation report (SITREP) to someone else. This person will then decide what to do next. It is anticipated that the unmanned system will be autonomous which is meant to further decrease the demands on the Soldier. Parameters, set in the system, will tell the robot how to move through the environment and what to do when it encounters an obstacle. One potential problem with automation is that once the operator chooses a path for the robot, it is assumed that the robot can get there easily; this may not always happen. For example, if the robotic asset encounters a small log, it may try various movements for several minutes and then signal the operator that it is in trouble. The
terrain is one factor that can impact how often the robot gets into trouble. A plan is entered in the robot, based on terrain data. If the asset gets in trouble, in order to determine the best path to take, the operator will have to retrace the robot’s steps to decide what to do. This will require time and operator resources. Further, communications are a high frequency event. Communications can come from various areas as well as be transmitted by operators to lower and upper command levels. The hypothetical information flow in figure 1 was derived from subject matter expert (SME) input. It shows the inherent complexity of the future robotic environment.

![Hypothetical information flow diagram](image)

Figure 1. Hypothetical information flow.

Thus, any major change in the future force, such as the addition of robotic systems, implies problems as well as solutions. Robotic systems with diverse roles, tasks, and operating requirements are being designed to exploit future battle spaces. The role of the human operator is not well understood; however, most of the contemplated systems will require active human control or supervision with the possibility of intervention. In the most extreme case, Soldiers will operate multiple systems while “on the move” and while undergoing enemy fire. In all cases, the workload and stress will be variable and unpredictable—changing rapidly as a function of the military environment.
The U.S. Army Research Laboratory’s (ARL’s) Human Research and Engineering Directorate has begun a large-scale research endeavor with the U.S. Army Tank and Automotive Research and Engineering Center (TARDEC) to develop a common user interface that maximizes multi-functional Soldier performance of primary mission tasks by minimizing required interactions and workload in the control of ground and air robotic systems. ARL is investigating various technologies and concepts to achieve this minimization, one of which is adaptive automation.

The purpose of this report is to investigate automation technologies that unload the war fighter interacting with unmanned systems during multi-tasking missions. First, we investigate automation technologies, specifically their positive and negative effects on human performance and situation awareness (SA). Next, we discuss adaptive and adaptable processes as methods that potentially overcome the disadvantages of pre-set automation. The last section discusses our research approach for assessing the effectiveness of implementing automation for control of robotic systems.

1.2 Automation

Future robotic systems are being designed to be used in many facets of the modern battle space and to the degree possible, to be autonomous. This requires rapid response capabilities and intelligence to be built into the system. However, ultimate responsibility for system outcomes always resides with the human, and in practice, even highly automated systems usually have some degree of human supervisory control (Woods, 1996). Particularly in combat, some oversight and the capability to override and control lethal systems will always be a human responsibility (Barnes, Wickens, & Smith, 2000). Human involvement is necessary for system safety reasons in order to capture any changes in the commander’s goals and implied meta-goals and to minimize fratricide.

1.2.1 Levels of Automation

Automation is not an all-or-nothing phenomenon. Automation can vary in the degree to which a particular function that was previously performed by a human operator is allocated to a machine agent. That is, level of automation (LOA) can be fully autonomous, in which the machine agent performs the entire task, to manual, in which the human operator does (Sheridan, 1992). Automation can vary in dimensions and the stage of human information processing to which the automation is applied (e.g., information acquisition and analysis). Parasuraman, Sheridan, and Wickens (2000) developed a taxonomy of human automation control which has two dimensions: degree of automation (i.e., control options) and type of information processing function. The four information processing functions in the automation taxonomy are information acquisition, information analysis, decision making or action selection, and action implementation, which are similar to the observe-orient-decide-act or OODA loop in the parlance of military command and control. The taxonomy also captures the multiplicity of control options from fully automated to fully manual for each of these functions. The decision space in the taxonomy is complex and the inherent complexity implies that there is no single solution to partitioning control between the
human operator and the machine agent. Specifically, as the type of task the operator performs changes, the control logic may need to change as well. The taxonomy does not consider any other tasks the operator may be performing or the overall workload and stress imposed by the current environment. The addition of tasks and increased workload may complicate the problem of how to partition control. A review of the human performance literature reinforces the notion that there is no single solution to partitioning control; human performance varies greatly, depending on the operator task and the current environment (e.g., Parasuraman & Riley, 1997; Parasuraman et al., 2000).

1.3 Human Performance Issues for Automated Systems

1.3.1 Trust in Automation

Numerous problems related to human performance in automated systems have been identified in the literature. One problem with automated systems is the operator’s trust and level of use of the automation. Parasuraman and Riley (1997) compiled research and real-world examples of automation misuse, disuse, and abuse. The human operator ignored important indicators, failed to use reliable systems, misused unreliable systems, or misunderstood the true state of the system.

1.3.1.1 Over-reliance

Mosier and Skitka (1996) examined cases of automation bias, that is, bias toward relying on the automation. For example, in one case of automation bias, the operator tended to over-rely on automated systems even in cases when appropriate operator intervention would have averted performance problems. Mosier and Skitka did not identify a generalized automation bias per se but rather identified a number of performance problems related to lowered vigilance, high workload, time stress, and loss of SA because of this bias.

1.3.1.2 Under-reliance

Research has also shown that although there were circumstances when humans over-relied on automation, there were other equally important instances when they should have relied on automation and did not. Research by Dzindolet and colleagues has shown opposite results (under-reliance). In a study similar to the Battle Command Information System (BCIS) study, college students were asked to decide whether a target was present in or absent from the display (Dzindolet, Beck, Pierce, & Dawe, 2001). After each trial, target advisories purported to be from an aid (automatic target recognition [ATR] device) or a “peer” were given to the students. Participants were told the relative accuracy rates for the aids (or peers) and their own decisions. Based on this information, they had to decide whether to base future decisions on their own performance or on that of the automated advisory (or peer). Surprisingly, even when the aid made half as many errors (and the participants’ reward depended on accuracy), 80% of the participants chose to make their own decisions. Participants trusted peer advisories more than the aid advisory with the same accuracy level. They rationalized their decisions in terms of self
reliance. The most salient difference between this study and the BCIS study was that participants were told the ATR-peer decision after they had made their own decision. Thus, there was no workload advantage to using the aid advisory because the operator’s decision was made before the aid results were known. This is important because a simple manipulation (the order in which information about system reliability was provided) caused an automation bias to shift to a self-reliance bias. The same results were repeated in subsequent experiments but the self-reliance bias was mitigated if the individual was informed of the reason for the ATR errors and was given appropriate feedback during the initial trials (Dzindolet et al., 2001). Anecdotal data also indicate that there is a mistrust of aids in cases when there are high false alarm rates (Parasuraman & Byrne, 2003). In summary, humans are neither universally over-reliant or under-reliant on automated systems. The crucial factors seem to be workload, time stress, false alarm rate, and decision order.

1.3.1.3 Effect of Task Type, Workload, and Error on Automation Reliance

Automation reliability has the same contradictory effects on performance, depending on the task, workload, and type of errors that the automated device makes. Some research has shown no effect of aid reliability on performance (e.g., Dzindolet et al., 2001; Parasuraman, Molloy, & Singh, 1993; Singh, Molloy, & Parasuraman, 1997). The most likely reason is the lack of consistency between these studies in the tasks given to the participants. For example, in the Dzindolet study, participants received only one reliability level (60%, 75%, or 90%) and they were not given the system information that may have allowed them to respond effectively. The literature suggests that humans have problems understanding probabilities and may need some form of intervention in order to perform efficiently (Barnes, 2003; Wickens & Holland, 2000). Reliability level effects depend on operator strategies and on the type of error the aid manifests. Meyer (2001) has shown that when automation reliability is such that malfunctions are almost always correctly indicated (i.e., the automation makes few misses) then the operator has high reliance on the automation. This is an effective strategy but can result in a problem when the automation does miss because of the complacency effect. On the other hand, if automation reliability is such that few false alarms are made, then the operator has high compliance; if an automated alarm sounds, then the operator tends to immediately comply with the alarm and tend to the situation. Reliance on automated aids permits the operator to attend to tasks other than the automated task until the alert is triggered, thus improving multi-task performance and not just the performance of the automated task. Research using realistic UAV operator tasks indicated that reliant behaviors are affected primarily by the misses. In contrast, compliance errors are affected by the false alarm rate but not the miss rate of the automated device. However, increasing the operator’s workload and decreasing the aid’s reliability level had adverse effects on compliance and reliance errors.

The issue of automation reliance is complicated because performance depends on the type of processing task the operator performs, and paradoxically, high system reliability can result in costs as well as benefits. For example, Rovira, McGarry, and Parasuraman (in press)
investigated automation of artillery targeting decisions. For their particular task, Rovira et al. showed that reliable automation improved the commander’s decision latency without sacrificing accuracy. However, particularly for decision tasks related to choosing a course of action, higher reliability hurt the operator’s performance when the aids gave incorrect information. The surrogate commanders trusted 80% accurate aids more than the 60% ones in cases when they should have been more skeptical (i.e., when the aids gave them incorrect information). Apparently, the advisories from “trusted” aids were not scrutinized as thoroughly as those from less reliable aids.

One interpretation of the previous results is to assume that reliable aids lulled the operator into a false sense of complacency (e.g., Horrey & Wickens, 2001). However, the complacency literature suggests that the results depend on other factors than trust in the automation. Further, several researchers had failed to show complacency effects. Parasuraman et al. (1993) investigated the possible reason for the lack of complacency effects in a multi-task aviation environment. They showed that complacency did not occur for low workload (single task) conditions. However, for the high workload task, the operator became complacent (over relied) on the aid when the aid had a constant reliability level; when the reliability level varied over a block of trials, the performance decrement was ameliorated. This suggests that complacency is not only a function of trust in the automation but also the strategy used by the operator to handle high workload. For example, if an automation aid acted in a predictable manner (constant reliability), then the operators would commit their resources to other tasks in a high workload environment which in turn would cause a performance decrement in the unmonitored automation task.

High workload is a major factor in the efficacy of automation. When workload is high, the operator may trade SA for decreasing his/her workload by depending on the automation aids, even in cases when it was not beneficial to do so. This is not universally true; in some cases, automation improved overall performance even when the automated task required intervention because the operator’s residual cognitive capacity was allocated effectively among the set of tasks (Galster & Parasuraman, 2003; Lorenz, Di Nocera, Rottger, & Parasuraman, 2002). Too often, the loss of SA related to inefficient automation monitoring leads not only to performance decrements but also to an increasingly impoverished understanding of the work environment, which can result in catastrophic errors over time (Endsley, 1996; Mosier & Skitka, 1996; Parasuraman & Riley, 1997). Because of the uncertainty and risk associated with military environments, the use of automation must be done with extreme caution. The Soldier and his chain of command need to maintain SA; keeping the Soldiers out of the loop will have consequences for the immediate task and will predispose them to miss important cues that signal change (Endsley, 1996). Conversely, requiring Soldiers to engage in multiple tasks could very well have the same consequences.
1.4 Additive Principles

A possible solution is to create enough opportunity for flexibility in the system to ensure more automation during peak workload and greater operator engagement during low workload. The conventional view is that, generally, the greater the obvious workload, the more likely that a detrimental effect will be seen on overall performance. A possible solution would be for the operator to decide when to automate and which tasks to automate as mission requirements change. However, having the operator decide automation implementation is not always practical; it would burden operators with additional tasks precisely when they are already heavily loaded. For this reason, a number of researchers have suggested using some form of behavioral indicator to change levels of automation dynamically as a function of the changing work environment, namely, adaptive automation (Byrne & Parasuraman, 1996; Parasuraman, Bahri, Deaton, Morrison, & Barnes, 1992; Rouse, 1977; Scerbo, 1996).

The adaptive automation concept was proposed by Rouse about 25 years ago (1977). However, the technologies needed for its effective implementation were not readily available until recently. In adaptive systems, the “division of labor” between human and machine agents is not fixed but dynamic, in contrast to systems where provision of computer aiding is pre-determined at the design stage and task allocation is fixed during system operations. Adaptive automation uses mitigation criteria that drive an invocation mechanism to maintain an effective mixture of operator engagement and automation for a dynamic multi-task environment (as shown in figure 2). The invocation mechanism is triggered by whatever measurement process is used to represent the current state of the operator and/or task. If properly instrumented, the results of the measurement process should be displayed to operators in order to keep them informed of the state of the invocation process.

![Figure 2. Example of a closed loop adaptation for A, automated; A/M, automated/manual; and M, manual task sets.](image-url)
The adaptive automation process is more complex than simply unloading (or engaging) the operator of a task. To be effective, the invocation process must be sensitive to the operator’s combined tasking environment, which depends on interactions among tasks as well as overall workload, stress and safety considerations (Wickens & Holland, 2000). For example, the system might automate auditory tasks when the communication traffic reaches a predefined level but not change other task states until the overall workload measure (which could be a physiological index) reaches criterion (Dixon, Wickens, & Chang, 2004). Furthermore, whenever certain critical events occur, the invocation mechanism must be sensitive to indices that imply that the operator requires emergency automation (e.g., loss of consciousness because of g forces as evinced by physiological measurements) (Barnes & Grossman, 1985).

1.4.1 Characteristics of Adaptive Automation Systems

1.4.1.1 Invocation Methods

The method of invocation is a key issue in adaptive automation. Parasuraman et al. (1992) reviewed the major invocation techniques and divided them into five main categories:

- Critical events,
- Operator performance measurement,
- Operator physiological assessment,
- Operator modeling, and
- Hybrid methods.

The critical events method is exemplified by the work of Barnes and Grossman (1985). In this approach, automation is invoked only when certain tactical environmental events occur. For example, in an aircraft air defense system, the beginning of a “pop-up” weapon delivery sequence leads to the automation of all defensive measures of the aircraft. If the critical events do not occur, the automation is not invoked. Thus, this method is inherently flexible and adaptive because it can be tied to current tactics and doctrine during mission planning. This flexibility is limited by the fact that the contingencies and critical events are anticipated. A disadvantage of the method is its possible insensitivity to actual system and human operator performance. The critical events method will invoke automation, regardless whether the pilot requires it when the critical event occurs.

One potential way to overcome this limitation is to measure operator performance and/or physiological activity. In the operator performance measurement and operator physiological assessment method, the operator mental states (e.g., mental workload or more ambitiously, operator intentions) may be inferred from performance or other measures. The measures would then be used as input for the adaptive logic. For example, from performance and physiological measurements, it could be inferred that a human operator is dangerously fatigued or experiencing
extremely high workload. An adaptive system could use these measurements to provide computer support or advice to the operator that would mitigate the potential danger.

In the adaptive automation literature, Scerbo and colleagues (e.g., Prinzel, Freeman, Scerbo, Mikulka, & Pope, 2000) have conducted an elegant series of experiments that use electroencephalography (EEG) activity to drive adaptive automation. EEG is a non-invasive recording of the fluctuations in electrical activity of the brain. A general assumption is made that changes in EEG reflect arousal and workload (Scerbo, 2001). It is a closed loop adaptive system that moderates workload by decreasing the task demands when workload increases. Increases in workload are assessed via EEG. EEG is measured and an EEG engagement index is derived from alpha and beta components of the frequency domain. The system allocates the tasks based on the engagement index. A high engagement index is related to a high state of alertness and an increased ability to attend to stimuli (Freeman, Mikulka, Prinzel, & Pope, 1999). The Multi-attribute Task Battery (MAT) (Comstock & Arnegard, 1992) is used as the task test bed and includes a monitoring task, a resource management task, a communications task, and a tracking task. The tracking task shifts between manual and automated, depending on the engagement ratio of the operator. Tracking performance improved with an adaptive policy wherein high engagement EEG ratios invoked automation and low ratios invoked manual tracking compared to the opposite invocation policy (non-adaptive; switch to auto-low and manual-high). Prinzel et al. (2000) showed that when the engagement index increased and the system automated the task, performance of the tracking task was better than when the task was always in the manual condition.

However, these studies failed to show superiority of one invocation mechanism over another (e.g., adaptive automation versus simple automation of the tracking task). For example, a recent study by Mikulka, Scerbo, and Freeman (2002) used a similar paradigm to investigate vigilance performance during low, medium, and high event rates. They also introduced a control to ensure that the invocation policies based on the EEGs were the cause of the performance differences. Each EEG participant had a yoked partner who switched to automation (or manual) mode, based on task difficulty level (unrelated to his or her EEG state) as the EEG partner. The group that had an adaptive EEG policy was again superior to the non-adaptive EEG engagement policy group. However, the yoked groups followed the same pattern, which implied that the results were caused by task difficulty and switching policy and not the EEG indices per se. This suggests that modeling the task environment might be as effective as using the EEG engagement policy. The third invocation method is operator modeling. Human operator states and performance may be modeled theoretically, and the adaptive algorithm would be driven by the model parameters. Intelligent systems that incorporate human intent inferencing models have been proposed and are worth consideration (Geddes, 1985).

The final approach, hybrid methods, combines one or more of these different invocation techniques so that the relative merits of each method can be maximized in order to minimize operator workload and performance. It is possible that one measure or methods will not be able
to capture the complexity inherent in human performance. In a limited sense, hybrid methods of invoking adaptive automation have been tested. Wilson, Lambert, & Russell (2000) used a multiple measure approach to design a physiologically based adaptive automotive system. Wilson et al. had participants complete the MAT task with two levels of difficulty (varied the number of events that occurred in 5 minutes). EEG, electrocardiogram, electro-oculography, and respiration were measured during the task. Wilson et al. trained an artificial neural network (ANN) to recognize the physiological patterns that differentiate states of rest, low task difficulty, and high task difficulty. The ANN was then used to determine which condition a participant was performing, and when the high difficulty task was detected, the monitoring and auditory tasks were automated. Results showed that the ANN correctly identified the task conditions, and when adaptive automation was implemented, tracking error decreased and performance of the resource management task increased, compared to the manual condition. No comparison was made between fully and adaptively automated performances (Wilson et al., 2000).

1.4.2 Adaptive Versus Adaptable Automation

In adaptive systems, the decision to invoke automation or to return an automated task to the human operator is made by the system via any of the previously described invocation methods. This immediately raises the issue of user acceptance of such a system. Human operators may be unwilling to accede to the “authority” of a computer system that mandates when and what type of automation is or is not to be used. Apart from user acceptance, the issue of system unpredictability and its consequences for operator performance may also be a problem. It is possible that the automated systems that were designed to reduce workload may actually increase it. Billings and Woods (1994) cautioned that a truly adaptive system may be problematic because the system’s behavior may not be predictable to the user. To the extent that automation can hinder the operator’s SA by taking him or her out of the loop, unpredictably invoked automation by an adaptive system may further impair the user’s SA. However, if the automation were explicitly invoked by the user, then presumably, the unpredictability will be lessened. However, involving the human operator in making decisions about when and what to automate can increase workload. Thus, there is a trade-off between increased unpredictability versus increased workload in systems in which automation is invoked by the system or by the user. Opperman (1994) characterized these alternatives as “adaptive” and “adaptable” approaches to system design (see also Scerbo, 2001). In either case, the human and machine systems adapt to various contexts, but in adaptive systems, automation determines and executes the necessary adaptations. In adaptable systems, however, the operator is in charge of the desired adaptations. The distinction is primarily one of authority. In an adaptable system, the human always maintains authority to invoke or change the automation, whereas this authority is shared in an adaptive system. Inagaki’s (1999) design concept of “situation-adaptive autonomy” is related to this view of an adaptive system, but in his approach, control of a process is exchanged between human and computer in real time, based on time criticality and the expected costs of human and machine performance.
Although in this review we primarily consider how adaptive automation affects system performance, it is important to keep in mind that adaptable automation may provide an alternate approach with its own benefits. The level of automation (LOA) concept introduced by Sheridan (1992) does not specify which level should be used or who decides that there should be a change in level. When the decision is made by a designer before system operation, it is a part of system design and corresponds to picking an appropriate LOA for that system design. The decision can also be made by automation itself (or some expert system) during system operations as a part of a truly adaptive automation system. In both of these cases, the human operator is not involved in the decision. In adaptable systems, however, the human operator is involved in the decision of what to automate, similar to the role of a supervisor of a human team who delegates tasks to team members, but in this case, tasks are delegated to automation. The challenge for developing such an adaptable automation system is that the operator should be able to make decisions regarding the use of automation in a way that does not create such high workload that any potential benefits of delegation are lost.

1.4.3 Human Interaction With Adaptive Systems

Since the theoretical frameworks for adaptive automation proposed by Rouse (1988) and Parasuraman et al. (1992), there has been a steady stream of empirical work aimed at examining the effects of adaptive automation on human and system performance in different application domains (e.g., Hilburn, Jorna, Byrne, & Parasuraman, 1997; Kaber & Riley, 1999). The initial studies were designed to investigate whether the performance costs of certain forms of static automation (described previously) such as reduced SA, complacency, and skill degradation can be mitigated by adaptive automation. Most of these studies used a critical event or model-based approach for adaptive automation. A task was allocated dynamically to human or machine control at some time during a simulated mission when some critical event occurred or as dictated by a simple model of operator and system performance. For example, Hilburn and colleagues (1997) examined the effects of adaptive automation on the performance of military air traffic controllers who were provided with a decision aid for determining optimal descent trajectories of aircraft—a descent advisor (DA). The DA was present at all times (static automation) or only when the traffic density exceeded a threshold. Hilburn et al. found significant benefits for controller workload (as assessed with the use of pupillometric and heart rate variability measures) when the DA was provided adaptively during high traffic loads, compared to when it was available throughout (static automation) or only at low traffic loads. In addition to physiological measures of workload, other measures can be used to assess the workload-leveling effect of adaptive automation. Kaber and Riley (1999), for example, used a secondary task measurement technique to assess operator workload in a target acquisition task. They found that adaptive computer aiding based on the secondary task measure enhanced performance of the primary task.
The results of these and other studies (see Parasuraman, 2000, for a review) indicate that adaptive automation can reduce the problem of imbalanced workload, with attendant high peaks and troughs, which static automation often induces. As discussed previously, during high workload, operators tend to adopt an attention allocation strategy that results in diminished monitoring of an automated task (Parasuraman et al., 1993; Moray, Inagaki, & Itoh, 2000). As a result, operators can miss malfunctions in the task or fail to correct suboptimal performance by the automation because they are busy attending to other tasks. Adaptive automation in the form of a temporary return of the automated task to human control can mitigate this so-called complacency effect. In a study with the MAT flight simulation battery, Parasuraman et al. (1999) showed that temporary return of an automated engine systems task to human control benefited subsequent operator monitoring of the task when it was returned to automated control. It is important to emphasize that the reallocation to human control was brief. If the benefit could only be obtained by prolonged human intervention in the task, that would defeat the purpose of automating the task in the first place. Parasuraman et al. found that the benefit of adaptive reallocation was found for either of two methods of invocation, a model-based approach in which the temporary return to human control was initiated at a particular time specified by the model, or a performance measurement approach in which the adaptive change was triggered only when the operator’s performance of the engine systems task fell below a specified level. A subsequent study showed that the operator (and system) performance benefit could also be sustained for long periods of time, in principle, indefinitely, by repetitive or multiple adaptive task allocation at periodic intervals (Mouloua, Molloy, & Parasuraman, 1993). Such brief, periodic, adaptive reallocation of an automated task to human control can enhance overall system performance by maintaining the operator’s awareness of the automated task parameters or by refreshing the operator’s memory (his or her “mental model”) of the automated task behavior. In support of the latter explanation, Farrell and Lewandowsky (2000) showed that they could successfully computationally model the complacency effect and the benefit of adaptive reallocation in a three-layer connectionist network with a memory decay function for nodes representing automation performance.

These results show that adaptive automation can balance operator workload and reduce automation complacency. However, Parasuraman et al. (1999) also showed that performance benefits can be eliminated if adaptive automation is implemented in a clumsy manner, which supports the concerns of Billings and Woods (1994). Moreover, these studies, while clearly pointing to the potential benefit of adaptive automation, had some limitations. The model-based invocation method used in many of the studies has the advantage that the model can be implemented off line and easily incorporated into rule-based expert systems. However, this method requires a valid model, and many models may be required to address all aspects of human operator performance in complex task environments.
2. Research Approach

2.1 Robotic NCO (Noncommissioned Officer) Program

Automation technologies have been applied to basic multiple task (i.e., the MAT battery) and aviation environments. Research to date has not examined the feasibility of adaptive or adaptable automation in an environment in which an operator will control multiple robotic systems from a single control unit. The robotic environment for the Future Force will be highly complex. In order to control the task types and levels of automation, a new simulation test bed was developed by ARL in collaboration with George Mason University, which emulates the essential robotic tasks. The program, robotic NCO, was based on an existing prototype operator control units (OCUs) designed by Micro Analysis & Design and TARDEC. Figure 3 shows a picture of this OCU.

![Prototype operator control unit.](image)

2.2 Program Specifications

Robotic NCO is a multi-task environment that includes three main tasks: unmanned ground vehicle (UGV) control, UAV sensor use, and multi-level communications. The program runs on Java 1.5 and uses timers to initiate the tasks. The overall speed of the UGV and UAV and duration of the scenario is manipulated by the experimenter. The scenarios are created by the experimenter in the script creator (see figure 4). First, the paths for the UAV and UGV are drawn with the mouse and then the units (enemy or friendly) are selected from a menu and placed on a map that depicts the common operational picture (COP). This map is intended to support SA. The experimenter can also program the locations at which the UGV will stop. The communications are scripted on the basis of the time they will be presented. The specific communications that will be received by the operator are typed into the script.
3. Experiment I

3.1 Participants

Participants were 17 (10 women, 7 men) students from George Mason University (GMU). The mean age of the group was 22 (19 to 26 years old). Participation in the experiment was strictly voluntary and participants could withdraw from the experiment at any time. The experiment lasted approximately 2 hours and participants were paid $15.00 per hour.

3.2 Apparatus

3.2.1 Robotic NCO

Robotic NCO required participants to complete three tasks simultaneously: (a) respond to targets encountered by a UAV by using the mouse; (b) respond to potential obstacles and waypoints encountered by a UGV by using the mouse; and (c) respond to audio communications via the keyboard. These communications consisted of call sign acknowledgments and answering “yes/no” SA questions. The program was run with a Dell PC, keyboard, mouse, and monitor.
3.2.2 Workload and Situation Awareness Questionnaires

A subjective overall rating was given at the end of each trial on the participant’s perceived overall workload (OW) and SA. This was a single number from 0 to 100 for each rating. OW and SA criteria were adapted from the National Aeronautics and Space Administration Task Load Index (NASA-TLX) (Hart & Staveland, 1988) and the Cognitive Compatibility Situation Awareness Technique questionnaire (CC-SART) (Taylor, 1990), which were both administered in their complete form at the end of all 16 trials.

3.3 Experimental Scenario

The general scenario was as follows: one UGV and one UAV were supervised by a single operator for a mission. The UAV and UGV starting point, ending point, and path were scripted before the operator use. The UAV traveled faster than the UGV and provided surveillance/reconnaissance information to the operator. The UGV followed its routed path and when an event occurred, waited for operator input. While supervising the two robotic units, the operator received communications.

More specifically, in the first experiment, participants were asked to take the role of a robotic operator in a Mounted Combat System (MCS) company. The participant conducted a reconnaissance mission for the MCS platoon. To complete the mission, the participant used two robotic systems: a UGV and a UAV. Intelligence identified areas of interest (designated on the COP map with letters) and way points had been planned for the UGV to look at these named areas of interest (AOIs). The UAV had also been waypoint planned to view around the areas. During the reconnaissance mission, the UAV received electronic intelligence (ELINT) hits from possible targets which were displayed in the UAV view as white squares. When the participant saw an ELINT marker, s/he focused on that location and identified the possible target as enemy or friendly. The target was then displayed on the map.

At the same time, the UGV moved through the area via pre-planned waypoints. During the mission, the UGV stopped and the UGV status bar flashed. When the participant noticed that the UGV had stopped, s/he clicked on the UGV bar which accessed views (simulating video images) from the UGV, a picture of the UGV obstacle. The UGV view was displayed in the same location as the UAV view on the computer monitor. There were two reasons why the UGV stopped; it reached the named AOI or encountered an unknown obstacle. When the UGV reached the named AOIs, the participant conducted reconnaissance of the area and then caused the UGV to resume along its pre-planned path by selecting “continue”. When the UGV encountered an unknown obstacle, it was one of two types, a blocking obstacle (e.g., log, ditch) or a traversable obstacle. A blocking obstacle required re-planning and the participant re-routed the UGV by selecting “re-route”. A traversable obstacle required the participant to cause the UGV to resume along its pre-planned path by selecting continue. The design of the UGV manipulation was based on field observations of a UGV. The current autonomous navigation system operates in the manner used in this experiment.
Communications were also received during the mission. There were two types of communications: call sign acknowledgment and status reports. The participant heard call signs at random intervals, which were ignored or acknowledged, and at various times, s/he was prompted for UGV or UAV status update and location of targets. These status communications were designed to assess SA. Figure 5 shows screen shots of what the participants saw on the robotic NCO simulation during their mission.

**Zoomed In UAV View**

- UAV Status Menu
- Possible Target
- MAP with UAV and UGV Path
- Some Friendly Units

**Non-Zoomed in UAV View**

- UnZoomed UAV View
- Elint Marker Present
- MAP with UAV and UGV Path
- Enemy, Friendly, Unknown Units

Figure 5. Robotic NCO screen shots.

At the end of the mission, the participant was given a map with two paths drawn on it. Based on all the information retrieved from the UAV and UGV, the participant chose the safest path for
the MCS platoon to take. The participant was instructed to choose the path that would minimize the exposure of the platoon to enemies in the area.

3.4 Procedure

Upon arrival at the lab at GMU, the experimenter briefed the participants about the purpose of the experiment to allow them to ask questions about the study and to determine if they would like to participate. During this briefing, participants were told the study’s purpose and general procedures. That is, a description of the NCO program was given, and the primary goals of the experiment were briefly outlined. They were told that participation was *strictly voluntary* and that they could withdraw from the experiment at any time.

The program robotic NCO was described to the participants, and the functionality of each button on the program was explained. The participants were then put through one practice mission to ensure that they understood the task. After familiarization, the participants complete 16 missions.

3.5 Experimental Design

The design of the experiment was a 2 x 2 x 2 within-subjects design. The within-subjects factors were the number of targets to be identified in the UAV task (UAV; low, 10 possible targets, high, 20 possible targets), number of times the UGV requested operator assistance (UGV; low, three AOIs, two obstacles, high three AOIs, four obstacles) and the priority of the incoming communications (COMMS; low, 16 low priority and four high priority messages; high, four low priority and 16 high priority messages).

More specifically, each task was manipulated to produce a “high” and “low” condition in the following ways.

3.5.1 UGV

For the high condition, the UGV stopped seven times (three AOIs and four UGV obstacles) in which the participant had to attend to the UGV by selecting the UGV status bar. The UGV would not move forward until each AOI or obstacle was identified by the operator. For the low condition, the UGV stopped five times (three AOIs and two UGV obstacles) in which the participant had to attend to the UGV as described before.

3.5.2 UAV

For the high condition, the participant was presented with 20 targets in the UAV view. The participant clicked the ELINT marker (i.e., white squares) with the mouse to obtain a clear image of the target. S/he then used a right mouse click to identify the target as enemy or friendly. An “enemy” was indicated by a red triangle, and an “unknown” was indicated by a yellow triangle. This procedure had to be completed within 6 seconds or the target would disappear, resulting in a
“missed” target by the participant. For the low condition, the participant was presented 10 targets and identified the targets as described before.

### 3.5.3 Communications

For the high condition, the operator was presented with 16 call signs to acknowledge (of 20). Call sign acknowledgment had to be performed within 5 seconds of the cue presentation or a “missed” call sign would result. For the low condition, the operator was presented with four call signs to acknowledge (of 20) during the same conditions as before. In each condition, the remaining four communications events were SA questions. There were two questions that assessed Type 1 SA (perception), two questions for Type 2 SA (comprehension), and one question for Type 3 SA (prediction).

The following SA questions were presented during the mission and required a “yes/no” response (Types 1, 2, and 3 refer to Endsley’s [1996] three-level taxonomy of SA):

- **Type 1 SA Questions:**
  1. Has the UGV required re-routing?
  2. Are there enemy troops in sector X?

- **Type 2 SA Questions:**
  3. Is the UAV in Sector X?
  4. Are there enemy troops within two grid squares of sector X?

- **Type 3 SA Question:**
  5. Is it likely that there are enemies in sector X? (This would pertain to a sector not yet encountered by the UAV, thus requiring prediction by the participant.)

The task manipulations of high and low were verified in a pilot study. In this experiment, participants completed 16 missions (two of each type). Each mission lasted 5 minutes. The order of the missions was counterbalanced.

The following dependent variables were measured during the experiment:

- **UGV Route Planning:** Reaction time to implement new route on blocking obstacle event or to observe an AOI and continue UGV.
- **Target acquisition:** Accuracy (proportion correct) and reaction time to identify target as friendly or unknown.
- **Communications:** Reaction time and percent missed for status responses.
- **Situation Awareness:** Overall SA as measured by the CC-SART, and mean number of correct responses on the SA questions presented as communications.
- **Overall workload**
4. Results

To examine the effects of the three task types (UGV, UAV, and communications) and task load manipulations (low and high) on performance, multivariate analyses of variance (MANOVAs), and subsequent ANOVAs were conducted.

4.1 UGV Task Performance—Reaction Time (RT) for AOIs and Obstacles

Figure 6 is a graph of mean RTs to UGV obstacles. Results for the UGV task showed that when the UGV stopped infrequently, participants were faster to respond to UGV obstacles when there were many UAV targets to identify than when there were a few. This pattern was not significant when the UGV stopped frequently.

A MANOVA revealed significant effects of UAV Targets x UGV Requests on UGV task performance and a main effect of UAV targets, $F(2,14) = 7.29, p = 0.03$ and $F(2,15) = 2.14, p = 0.00$, respectively. No other effects were significant. To determine whether RTs for AOIs or obstacles or both contributed to the significant interaction, ANOVAs were run. ANOVAs showed a significant effect for UAV Targets x UGV Requests for RT for obstacles, $F(1,15) = 17.96, p = 0.00$, but not for AOIs, $F < 1.0$. To explain this interaction, paired comparisons were conducted. Results showed that when the UGV task was high, there was no effect of UAV condition on RT for obstacle ($t < 1$). For the low UGV task, however, RT was significantly longer for the low UAV condition, $t(15) = 4.11 p = 0.00$, as shown in figure 7.

![Figure 6. Mean reaction times (seconds) for obstacles in the UAV and UGV conditions.](image-url)
4.2 UAV Task Performance—RT to Targets and Percent Correctly Identified

A MANOVA revealed a marginal effect of UAV targets on UAV task performance, $F(2, 15) = 3.37, p = 0.06$. No other effects were significant. To determine whether RTs to targets or the percent correctly identified or both contributed to the significant interaction, ANOVAs were run. ANOVAs showed a significant effect for UAV targets for the percent of targets correctly identified, $F(1,16) = 4.02, p = 0.05$, but not for RTs, $p > 0.10$. More targets were identified in the high UAV condition ($x = 94\%$) than the low UAV condition ($x = 92\%$).

4.3 COMM Task Performance—Percent of Missed Call Signs and RT to Communications

Figures 7 and 8 are graphs of mean RTs to communications. Results for the communications task showed that participants generally took longer to respond to communications when they had to also identify many UAV targets. Additionally, when the amount of high priority communications was low, participants took longer to respond to communications when they had many UAV targets and UGV requests.

![Figure 7](image1.png)

Figure 7. Mean RT (seconds) for communication events in the low COMM condition.

![Figure 8](image2.png)

Figure 8. Mean reaction time (seconds) for communication events in the high COMM condition.
A MANOVA revealed a marginal effect of UAV Targets x UGV Requests on communications task performance, $F(2,15) = 5.11, p = 0.02$. Other significant effects were, UGV Requests, $F(2,15) = 4.07, p = 0.03$ and COMMS, $F(2,15) = 5.71, p = 0.01$. To determine whether the percent of missed communications or the RT to communications or both contributed to the interaction, ANOVAs were run. ANOVAs showed a significant effect for UAV Targets x UGV Requests x COMMS for reaction time, $F(1,16) = 5.73, p = 0.02$, but not for missed communications, $F < 1.0$. To explain this interaction, separate ANOVAs were conducted for each level of COMMS. For the low COMMS condition, there was a significant effect of UAV Targets x UGV Requests, $F(1, 16) = 7.88, p = 0.01$. No other effects were significant. The interaction for high COMMS was not significant, $F < 1.0$. Paired comparisons for the low COMMS conditions were conducted. Results showed that for the low UGV condition, there was no effect of UAV difficulty on communications RT ($t < 1$). For the high UGV condition, however, communications RT was significantly longer for the high UAV condition, $t(16) = 3.40, p = 0.00$, as shown in figure 8.

### 4.4 Self-Reported SA Ratings

#### 4.4.1 Overall Situation Awareness

Figures 9 and 10 are graphs of mean SA reported at the end of each scenario. Participants in the low UAV task condition (fewer possible targets) reported better SA when the UGV had more stops and communications were more frequent than when the UGV stopped infrequently. A MANOVA revealed a significant effect for UAV Targets x UGV Requests x COMMS on SA ratings, $F(1,15) = 12.91, p = 0.00$, as well as a significant effect for UAV Targets x UGV Requests, $F(1,15) = 11.87, p = 0.00$, and UGV Requests x COMMS Requests, $F(1,15) = 4.38, p = 0.054$. No other effects were significant. To explain the UAV Targets x UGV Requests x COMMS Requests interaction, separate ANOVAs were conducted for each level of UAV target condition. For the low UAV condition, there was a significant effect of UGV Requests x COMMS requests, $F(1,15) = 11.66, p = 0.00$. The interaction for high UAV condition was not significant, $F < 1.0$. Paired comparisons for the low UAV condition were conducted. Results showed that when the communications task was low, there was no effect of UGV condition on overall SA, $t(15) = -1.56, p = 0.13$. For the high communications condition, however, SA was significantly higher for the high UGV condition, $t(15) = -3.98, p = 0.00$, as shown in figure 9.
4.4.2 Levels of Situation Awareness

4.4.2.1 Type 1 Situation Awareness

Figure 11 is a graph for type 1 SA questions presented during each scenario. Analyses showed that there was a significant interaction of UAV Target x COMMS for Type 1 SA, $F(1,16) = 5.68$, $p = 0.03$. The main effects for UAV Target and COMMS were not significant, $p > 0.10$, but was for UGV requests, $F(1,16) = 9.79$, $p = 0.00$. Perception was significantly better when there were fewer UGV stops ($x = 0.830$) than when the UGV made many stops ($x = 0.724$). To explain the
interaction of UAV Target x COMMS, paired comparisons were conducted. Results for the low COMMS condition showed that SA was higher when UAV requests were low relative to high, $t(16) = 2.48, p = 0.02$. There was no significant difference for the high COMMS condition, $p > 0.10$.

![Figure 11. Mean correct for type 1 SA communications in the UAV and COMMS conditions.](image)

### 4.4.2.2 Type 2 Situation Awareness

Analyses showed that there was a significant main effect of COMMS for Type 2 SA, $F(1,16) = 5.12, p = 0.03$. Comprehension was significantly better when there were fewer incoming high priority communications ($\bar{x} = 0.691$) than when there were many high priority communications ($\bar{x} = 0.577$). No additional interactions or main effects for UAV Targets, UGV Requests, or COMMS were significant.

### 4.4.2.3 Type 3 Situation Awareness

Analyses showed that there was a similar trend of COMMS for Type 3 SA as there was for Type 2 SA, $F(1,16) = 3.38, p = 0.08$. Prediction was better when there were fewer incoming high priority communications ($\bar{x} = 0.661$) than when there were many high priority communications ($\bar{x} = 0.566$). No interactions or main effects for UAV Targets, UGV Requests, or COMMS reached statistical significance.

### 4.5 Overall Workload

No interactions or main effects for UAV Targets, UGV Requests, or COMMS on overall workload reached statistical significance. Reported workload was relatively low. The mean workload, regardless of condition, was 39.6 (31.06 to 44.37).
5. Conclusions

5.1 Findings from Experiment I and Future Experimentation

Before the implementation of any automation scheme, whether it is adaptive or adaptable, we need an understanding of the robotic operators’ tasks and their abilities (or lack of) to complete them. The goal of the first experiment was to identify tasks in the human-robot interaction environment that were challenging to the operator (i.e., high driver tasks). We used the simulation environment, robotic NCO, to achieve this goal. In experiment I, operators were required to use the robotic systems (UAV and UGV) to identify enemy units in a pre-defined area and to respond to incoming communications. With the information received from the UAV and UGV, operators were asked to choose a safe path for a platoon to take through the reconnaissance area. Overall, the results showed that participants were good at integrating information received from the UAV and UGV to choose the platoon path. However, the data suggest that the multi-tasking requirements of the robotics NCO simulation diminished performance of the individual tasks. In general, participants generally took longer to respond to communications when they had to also identify many UAV targets and UGV requests. Additionally, when the amount of high priority communications was low, participants took longer to respond to communications when they had many UAV targets and UGV requests (see figure 7). This result appears to indicate that high priority but infrequently occurring communications pose a particularly high monitoring load on the operator, as suggested by recent studies of vigilance and monitoring in semi-automated systems (Warm, Dember, & Hancock, 1996). Participants also took longer to respond to UGV obstacles when there were fewer UAV targets to identify (see figure 6). It is unclear at this point why this occurred. Additional results were that when there were fewer targets to be identified with UAV but the UGV stopped frequently and there were many communications, overall SA was higher (see figures 9 and 10). Further comprehension was better when high priority communications were low. In general, performance (e.g., RT) was degraded when task load was high (i.e., many UAV targets and UGV stops).

Research conducted in more realistic simulation environments has also shown that in general, using multiple robotic systems has a negative effect on performance (e.g., Rehfeld, Jentsch, Curtis, & Fin cannon, 2005; Chen, Durlach, Sloan, & Bowens, in process). Rehfeld had participants indirectly monitor a robot(s) via a video feed sent from the robotic vehicle and identify a target in an urban setting. In general, participants had a difficult time identifying targets with one UGV and when they were given two UGVs to complete the target identification, performance did not significantly increase. However, when two operators were given one UGV, they found nearly 200% more targets than a single operator for the most difficult conditions.
Similarly, Chen et al. showed that single operators were not efficient when they had multiple systems; having additional systems for target acquisition did not significantly improve performance. Chen also examined control modality for the UGV, whether teleoperated or way point controlled. The teleoperated condition was the least efficient condition for target acquisition. When operators were given a UAV and UGV, they relied on the UAV rather than an integrated strategy. These results, along with the results of experiment I, suggest that even if robotic systems have minimum control requirements (waypoint control), targeting is a difficult task for the operator. When other factors in the multi-tasking environment of robotic operators are taken into account, such as communications or the stress of combat, performance may be degraded even further. The potential degrading effect of high communications load on Type 2 SA of the battlefield environment was also noted in the current study.

The current research investigations with the robotic NCO simulation and the simulation research by Chen and colleagues (in process) have identified several tasks to consider for automation or aiding target identification and robotic control. In the next series of studies with robotic NCO, we will implement some of these automation strategies. To evaluate the effectiveness of the automation, we will compare performance with and without its implementation. For example, with the UGV task, we will increase operator involvement with the UGV and compare manual operation to fully autonomous operation. For the UAV, we plan to investigate the effects of an automated target recognition system on operator performance. The ensuing studies will compare the effects of various types of automation (adaptive, adaptable, full, manual control) on Soldier performance while we vary the multi-tasking environment, including the reliability of the proposed aids.

### 5.2 Candidate Physiological Measures

In addition to the types of automation to be implemented, we are examining the best triggers for the automation. There are many ways in which the system changes can be initiated: subjective workload assessments, operator performance, and physiological measures. In a separate series of experiments, we will be assessing the utility of various physiological measures for invoking automation. Physiological measures can provide additional information that can be tapped for control of adaptive systems. When physiological measurement is used, the emphasis is on the operator’s capabilities, not the system’s. Measurement technology is developing rapidly and is showing improvements in the areas of non-intrusiveness, precision, and prediction. Therefore, it is important to evaluate these methods for their practicality and sensitivity to operator workload.

Kerrick, Oie, & McDowell (research in progress) are evaluating EEG in realistic motion environments. In the adaptive automation literature, a general assumption is made that changes in EEG reflect arousal and workload (Scerbo, 2001). Therefore, it is only logical that we consider EEG as a potential trigger for automation. One limitation of most physiological measurements, such as EEG, is their sensitivity to noise in the environment. Kerick and colleagues are assessing the feasibility of extracting high-quality, task-relevant EEG signals from
participants performing Soldier tasks. EEG measures (i.e., P300) will be used to examine if state-of-the-art measurement technologies and signal-processing techniques will allow the recording of relevant signals in conditions that will induce different levels of self-induced (e.g., muscular activity) and external (e.g., electromagnetic) noise artifacts. The ability to assess brain functions underlying performance in tasks and environments consistent with common Army operational scenarios is critical to the development of the ability to monitor Soldiers in real time—a key issue in using physiology for adaptive automation.

In a related effort, Cosenzo and Fatkin (research in progress) are evaluating cerebral blood flow velocity (BFV) changes in a multi-tasking simulation. BFV is assessed by trans-cranial Doppler sonography (TCD). TCD is used to assess the moment-to-moment changes in BFV. TCD is being evaluated because it is robust in noisy environments. It is also a relatively new technique in cognitive psychology and has not been used in adaptive automation applications. Research results may show that BFV is a valid measure of an individual’s cognitive status and may be used as a driver in an adaptive system. Future automation technologies may be driven by the patterns of electrophysiological activity that underlie cognition, whether it is EEG, TCD, or multiple measurements. This work will be used to determine the feasibility of physiological activity as a driver for an automation system.

5.3 General Conclusions

The U.S. Army Future Force requires Soldiers to have multiple responsibilities in an information-rich environment. Soldiers will be responsible not only for operating and fighting within their own vehicle but also for some set of robotic vehicles tethered to them, both air and ground. The addition of robots will present challenges to the Soldier (Mitchell, 2005). The combination of robotic operational tasks with other traditional military tasks will create high workload. For example, Mitchell (2005) showed that the gunner (traditional military tasks-scan for targets) was overloaded when he had a UGV that required operator intervention and as a result, the gunner decreased his scanning.

Studies conducted to date have identified some of the major issues (e.g., Mitchell, 2005) and the preliminary results (e.g., Chen et al., in process; Rehfeld et al., 2005) indicate that adaptive automation may be a useful mitigation strategy to help offset the potential deleterious effects of high cognitive load on Army robotic operators in a multi-tasking environment. There are human performance advantages and disadvantages of battlefield automation. In implementing any automation, we must keep in mind issues relating to trust in automations. Research has shown that individuals over-rely and under-rely on automated systems, depending on the following factors: the decision order, operator overload, false alarm rate, and reliability of the system. In addition, poorly designed automation can result in operator complacency and the loss of SA. In designing the automation, we will consider its effects on operator workload and performance as well as operator complacency and reliance on the automation.
From the work with the robotic NCO simulation, we will identify the candidate tasks and define adaptive architecture concepts to mitigate the Soldier’s workload and improve overall performance in the environment, based on some subset of adaptive logics. Work by Parasuraman, Galster, Squire, Furukawa, & Miller (2005) has demonstrated the efficacy of adaptable automation for aiding human control of multiple simulated robots; this will also provide input into the design of adaptive architectures. We will also use the work of Kerick et al. and Cosenzo and Fatkin (research in progress) to determine if physiological measurements are practical triggers for this environment. We intend to develop increasingly realistic simulations as we understand the efficacy of adaptable or adaptive options in multi-tasking environments. The adaptive concepts will be first evaluated in the robotic NCO program and then in prototype robotic OCUs developed by the Army. The ultimate goal of the program is to allow future combat vehicle operators to conduct remote targeting with aerial and ground robotic systems in a multi-tasking, high stress environment while maintaining sufficient combat awareness to ensure their survival.
6. References


<table>
<thead>
<tr>
<th>No. of Copies</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DEFENSE TECHNICAL INFORMATION CTR (PDF ONLY)</td>
</tr>
<tr>
<td>1</td>
<td>US ARMY RSRCH DEV &amp; ENGRG CMD SYSTEMS OF SYSTEMS INTEGRATION AMSRD SS T 6000 6TH ST STE 100 FORT BELVOIR VA 22060-5608</td>
</tr>
<tr>
<td>1</td>
<td>INST FOR ADVNCD TCHNLGY THE UNIV OF TEXAS AT AUSTIN 3925 W BRAKER LN STE 400 AUSTIN TX 78759-5316</td>
</tr>
<tr>
<td>1</td>
<td>DIRECTOR US ARMY RESEARCH LAB IMNE ALC IMS 2800 POWDER MILL RD ADELPHI MD 20783-1197</td>
</tr>
<tr>
<td>1</td>
<td>DIRECTOR US ARMY RESEARCH LAB AMSRD ARL CS OK T 2800 POWDER MILL RD ADELPHI MD 20783-1197</td>
</tr>
<tr>
<td>2</td>
<td>DIRECTOR US ARMY RESEARCH LAB AMSRD ARL CS OK T 2800 POWDER MILL RD ADELPHI MD 20783-1197</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR M DR M STRUB 6359 WALKER LANE SUITE 100 ALEXANDRIA VA 22310</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MA J MARTIN MYER CENTER RM 2D311 FT MONMOUTH NJ 07703-5630</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MC A DAVIS 320 MANSCEN LOOP STE 166 FT LEONARD WOOD MO 65473-8929</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MD T COOK BLDG 5400 RM C242 REDSTONE ARSENAL AL 35898-7290</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR ME MS A MARES 5800 CARTER RD FT BLISS TX 79916-3802</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MI J MINNINGER BLDG 5400 RM C242 REDSTONE ARSENAL AL 35898-7290</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MM DR V RICE BLDG 4011 RM 217 1750 GREELEY RD FT SAM HOUSTON TX 78234-5094</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MG R SPINE BUILDING 333 PICATINNY ARSENAL NJ 07806-5000</td>
</tr>
<tr>
<td>1</td>
<td>ARL HRED ARMC FLD ELMT ATTN AMSRD ARL HR MH C BURNS BLDG 1467B ROOM 336 THIRD AVENUE FT KNOX KY 40121</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED AVNC FIELD ELEMENT ATTN AMSRD ARL HR MJ D DURBIN BLDG 4506 (DCD) RM 107 FT RUCKER AL 36362-5000</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MK MR J REINHART 10125 KINGMAN RD FT BELVOIR VA 22060-5828</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MV HQ USAOTC S MIDDLEBROOKS 91012 STATION AVE ROOM 111 FT HOOD TX 76544-5073</td>
</tr>
<tr>
<td>NO. OF COPIES</td>
<td>ORGANIZATION</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MY M BARNES 2520 HEALY AVE STE 1172 BLDG 51005 FT HUACHUCA AZ 85613-7069</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MP D UNGVARSKY BATTLE CMD BATTLE LAB 415 SHERMAN AVE UNIT 3 FT LEAVENWORTH KS 66027-2326</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MJK J HANSBERGER JFCOM JOINT EXPERIMENTATION J9 JOINT FUTURES LAB 115 LAKEVIEW PARKWAY SUITE B SUFFOLK VA 23435</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MQ M R FLETCHER US ARMY SBCCOM NATICK SOLDIER CTR AMSRD NSC SS E BLDG 3 RM 341 NATICK MA 01760-5020</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MY DR J CHEN 12423 RESEARCH PARKWAY ORLANDO FL 32826</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MS MR C MANASCO SIGNAL TOWERS RM 303A FORT GORDON GA 30905-5233</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MU M SINGAPORE 6501 E 11 MILE RD MAIL STOP 284 BLDG 200A 2ND FL RM 2104 WARREN MI 48397-5000</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MF MR C HERNANDEZ BLDG 3040 RM 220 FORT SILL OK 73503-5600</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MW E REDDEN BLDG 4 ROOM 332 FT BENNING GA 31905-5400</td>
</tr>
<tr>
<td>1</td>
<td>ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MN R SPENCER DCSFDI HF HQ USASOC BLDG E2929 FORT BRAGG NC 28310-5000</td>
</tr>
<tr>
<td>1</td>
<td>ARMY G1 ATTN DAPE MR B KNAPP 300 ARMY PENTAGON ROOM 2C489 WASHINGTON DC 20310-0300</td>
</tr>
<tr>
<td>1</td>
<td>US ARMY RSCH LABORATORY ATTN AMSRD ARL CI OK (TECH LIB) BLDG 4600</td>
</tr>
<tr>
<td>1</td>
<td>US ARMY RSCH LABORATORY ATTN AMSRD ARL CI OK TP S FOPPIANO BLDG 459</td>
</tr>
<tr>
<td>1</td>
<td>US ARMY RSCH LABORATORY ATTN AMSRD ARL HR MR F PARAGALLO BLDG 459</td>
</tr>
<tr>
<td>1</td>
<td>US ARMY RSCH LABORATORY ATTN AMSRD ARL HR SE L PIERCE BLDG 459</td>
</tr>
<tr>
<td>1</td>
<td>US ARMY RSCH LABORATORY ATTN AMSRD ARL HR SE D HEADLEY BLDG 459</td>
</tr>
<tr>
<td>1</td>
<td>US ARMY RSCH LABORATORY ATTN AMSRD ARL HR SE K COSENZO (2 CYS) L FATKIN T BRANSCOME BLDG 459</td>
</tr>
<tr>
<td>4</td>
<td>US ARMY RSCH LABORATORY ATTN AMSRD ARL HR SE</td>
</tr>
</tbody>
</table>

34