Abstract

CLAMANN, MICHAEL PETER. The Effects of Levels of Invocation Authority on Adaptive Automation of Various Stages of Human Information Processing. (Under the direction of David B. Kaber.)

Adaptive automation (AA) has been explored as a solution to problems associated with extended periods of manual control and static automation in complex systems including fatigue, high workload, and loss of situation awareness. Several important decisions must be made regarding the implementation of AA in the design process, including who will decide when to invoke automation and what function or functions will be automated. Several studies have recently proposed models of automation based on four stages of human-machine system information processing, including information acquisition, information analysis, decision making, and action implementation. It is possible to apply AA to several of these functions simultaneously or independently. Management authority over automation (invocation authority) can be extended to a human operator, to a computer controller, or distributed between both through mutual suggestions and approvals of each authority. Both invocation authority over dynamic function allocations (DFAs) and the type of AA have both been shown to independently influence task performance, situation awareness, and workload in adaptive systems. It is possible that different levels of authority may have different effects on overall human-machine system performance when applied to various automated functions, such as detection tasks as compared to tasks requiring higher cognitive functions, like decision making.

The goal of this study was to assess the performance and workload effects of applying AA to four stages of human-machine system information processing using a performance-based approach and facilitating DFAs through two levels of computer authority (suggestion and mandate). The research was expected to provide insight into the existence of any interaction between these aspects of AA design. It was expected that higher level automation, such as information analysis and decision making, would be more compatible with mandated allocations, while lower levels, such as information acquisition and action implementation, would be more effective under partial human control. The additional task load imposed by the requirement of operator acceptance of DFAs suggested by a computer authority was expected to adversely affect performance of the more cognitive tasks.

Forty naïve subjects performed an air traffic control task as part of a dual-task scenario in which a secondary, gauge-monitoring task served as the objective measure of workload that controlled DFAs in the primary task. Each subject experienced one of four forms of automation (or no automation for control subjects) and both types of authority (suggest and mandate) during two trials (Each trial incorporated only one of the two authority types.). Results confirmed performance differences due to AA across the various aspects of information processing as well as between the two types of invocation authority. Specifically, subjects performed significantly better in the primary task during periods of automation as part of information acquisition AA as compared to decision making. During those same automated periods, subjects also performed significantly better when automation was suggested as compared to mandated. Contrary to the central thesis of the work that there would be a negative performance effect when computer suggestions were combined with AA of higher cognitive functions (information analysis and decision making), there was no evidence of an interaction effect of the two experimental manipulations. However, the individual results regarding automation type and authority were consistent with results appearing in the literature.

The results of this study provide evidence that the effectiveness of AA is dependent upon both the type of automation presented to an operator and the type of invocation authority designed into the system. This could potentially provide additional insight for effective AA design in complex systems.

THE EFFECTS OF LEVELS OF INVOCATION AUTHORITY ON ADAPTIVE AUTOMATION OF VARIOUS STAGES OF HUMAN INFORMATION PROCESSING

by
MICHAEL CLAMANN

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

PSYCHOLOGY

Raleigh

2002

Approved By:

_______________________________ _______________________________

____
Chair of Advisory

Biography

Michael Clamann abandoned the rocky mountains of Colorado to pursue a graduate degree in Ergonomics at North Carolina State University in August 2000. Michael grew up in Richmond, Virginia and attended James Madison University in Harrisonburg, Virginia where he received a dual degree in International Business and German. Prior to attending NCSU, he spent five years as a consultant specializing in the design and deployment of custom business software. His research interests include human-automation interaction and aviation human factors.

Acknowledgements

There are several people I need to recognize for their contributions to this project.

My advisor, David Kaber, provided guidance, support and access to his wealth of knowledge and experience.

Thanks to Lance Prinzel of NASA Langley Research Center, I had the opportunity to take part in the NASA Graduate Student Researchers Program, which funded this research.

My committee members, Mike Wogalter and Eric Wiebe, offered additional advice and direction along the way.

My uncle, Henry Nuttle, helped me make the transition from Denver to Raleigh and provided an open door to me both on campus and off.

My fellow graduate students generously gave me their time, diverse skills and perspectives, and valuable friendships. In particular, I am grateful to Srikanth Jayanthi for his contributions.

Laura Wiencek offered patience, understanding and unconditional emotional and spiritual support. Thanks, babe.

Finally, and most importantly, a very special thanks goes to my parents, Professor H. Peter Clamann and Eva Clamann for being my source of inspiration. Your strength makes this a truly special occasion.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AA	Adaptive Automation
AFA	Adaptive Function Allocation
ANOVA	Analysis of Variance
ATC	Air Traffic Control
DFA	Dynamic Function Allocation
FA	Function Allocation
GCAS	Ground Collision Avoidance System
MAT	Multi-Attribute Task
MCH	Modified Cooper-Harper
SA	Situation Awareness
SD	Standard Deviation
VAS	Visual Analog Scale

1 Introduction

Automation is pervasive in modern systems, including industries such as aviation, transportation, and retail. It is represented by autopilot systems designed to reduce pilot workload in extended flights, controls in automobiles that increase comfort by regulating interior temperature and maintaining consistent highway speeds, and computer systems that allow consumers to make purchases using their computers from remote locations outside business hours. Billings (1997) describes automation as a tool that allows a human operator to accomplish a task with less effort, or to complete a task that would have been impossible otherwise.

Wickens and Hollands (2000) offer four categories that summarize the reasons a designer might take advantage of automation when designing a system. The first includes tasks that cannot be performed by a human operator due to human limitations, such as the complex and dynamic calculations required for guiding a rocket, or manipulation of materials in hazardous environments. Another reason for automation is to assist with tasks that can be performed by human operators but result in poor performance or high workload for the operator. Autopilots, airborne collision advisory systems, and nuclear process control can be included in this category. Third, automation can augment, rather than replace human performance. In this category, automation serves as an aid rather than a replacement for the human operator. This is similar to the second category, except peripheral tasks are aided so the operator can focus on a central task. Examples include visual scratch pads that augment working memory by displaying computer input by keyboard or voice command. Finally, in many cases automation can be less expensive than human operators. Examples of automation of this type include robot workers in manufacturing plants, automated bank tellers, and telephone operators. Unfortunately, this final category tempts designers and managers to automate anything that has a monetary benefit, ignoring the operator who is left behind to monitor the system (Parasuraman and Riley, 1997). There are many important issues, such as determining what system functions should be automated and implementing alternatives to traditional static automation, that must be examined when designing an automated system. Each variable has its own unique advantages and disadvantages, which must be understood prior to designing an effective automated system. Several of these issues that can be controlled by designers are discussed in the following sections.

1.1 Adaptive Function Allocation

Although many modern systems have benefited from the introduction of automation, not all system functions must be allocated exclusively to machines. Responsibility can be divided between a human operator and an automated system. This division of tasks is referred to as function allocation (FA; Sanders and McCormick, 1993). The traditional approach to FA involves a firm division between the human operator and automation: a task is either automated all the time, or not at all. When designing using this approach to FA, task requirements are compared to the capabilities of humans and the capabilities of the automated system in order to determine which entity (human or automation) can most effectively perform the task. Tasks are then allocated either to the automation or to the human operator based on these comparisons (Sanders and McCormick, 1993). Scallen and Hancock (2001) cite several weaknesses in this traditional approach. For example, this approach can separate the human from the system, reducing the possibility for cooperation between the two. In addition, the extended periods of either manual or automated control brought about by an absolute division of tasks between human and automation can lead to new problems for the human operator. For example, manual control over extended periods often results in fatigue, increased error rates and reduced effectiveness. However, extended periods of automation can also have negative effects on a human operator. If a task is handed off to automation, the human's role becomes that of a monitor, a task for which humans are poorly equipped (Kaber and Endsley, in press). In these situations when the operator must remain in a control loop, but assumes a passive role in the system yielding to full automation, there are potential problems with the operator maintaining situation awareness (SA). Other results of extended automation can include complacency and skill decay (Parasuraman, 2000). In the event of an automation failure, any of these problems could lead to catastrophic outcomes.

Adaptive function allocation (AFA) provides a plausible solution to the challenges existing under traditional FA, such as managing workload, maintaining SA, preventing complacency, and maintaining operator skills. AFA differs from traditional FA in that task allocations may occur while the task is being performed. According to Scallen and Hancock (2001), in AFA, "the control of tasks shifts dynamically between humans and machines based on specified thresholds for environmental factors, operator confidence, or psycho-physiological factors." The form of automation that allows for these dynamic function allocations (DFAs) between automated and manual control is referred to as adaptive automation (AA; Kaber, Riley, Tan and Endsley, 2001; Hilburn, Molloy, Wong and Parasuraman, 1993). Adaptive automation systems react to situational demands by adjusting operating characteristics such as interface features or task functions. In these systems, the human operator and the machine assume a team relationship that facilitates optimal system performance (Bubb-Lewis and Scerbo, 1997). One goal of AA is to strike a balance between the potential excessive workload problems associated with manual control and the losses in SA resulting from extended periods of automated control (Wickens and Hollands, 2000).

Adaptive automation technology is still maturing, but some prototypes have already been produced (Scerbo, 1996). Citing Scott (1999), Scerbo et al. (2001) provide a real-world example for an application of AA in an aviation system involving the new Ground Collision-Avoidance System (GCAS) designed by the United States Air Force, Lockheed Martin, NASA, and the Swedish Air Force, which is currently being tested in the F-16D. The GCAS is designed for the rare occasion when excessive workload or a loss of SA prevents a pilot from maintaining a safe altitude, and provides the pilots with a tool that can react to emergency situations faster than their own reflexes. Piloting a modern military aircraft is a highly complex task in which operators are subject to a huge number of stimuli from both the cockpit and the environment. During periods of high workload, such as complex maneuvering or combat situations, a pilot can run the risk of missing an important signal. The GCAS is intended to be used during these periods of intense workload to prevent fighter pilots from losing SA and crashing into terrain. This automated system processes internal, external, and pilot input information to make in-flight calculations of altitude. If the system predicts that the aircraft will cross below a predetermined minimum altitude, the pilot receives a warning notifying him or her that automation is about to take over. If the pilot cannot adjust the aircraft's altitude heading accordingly within approximately five seconds of the warning, the GCAS assumes control. The task of adjusting the aircraft's heading is then offloaded to the automation, the aircraft's heading is corrected and the automation returns control to the pilot. The system provides appropriate feedback for the return to manual control so the pilot is informed of the system's status. This system is being evaluated for installation in three different fighter aircraft in the U.S. and Swedish Air Forces (Scerbo, Freeman, Mikulka, Parasuraman, & Prinzel, 2001).

When designing an AA system such as the GCAS, there are several important decisions that must be made regarding the implementation. These include who will make the decision to invoke automation, what function or functions will be automated, and the criteria to be used as a basis for triggering automation that determines the timing of invocations. These three issues are discussed in the following sections.

1.2 Locus of Control in Automated Systems

A critical issue that exists when designing a system to include AFA involves determining who makes the allocation decisions, also referred to as the locus of control (Parasuraman, 2000). In other words, should the computer or the human operator mandate the shift to, or from, automated control? Arguments exist for both operator-initiated and system-initiated invocation. Harris, Hancock and Arthur (1993) provide reasons in favor of operator-initiated invocations:

- Allowing the operator to maintain control over allocations helps keep him or her in the loop.

- Control over allocations provides the operator with a feeling of being in charge and can help maintain SA, both of which can help make workload more manageable.
- It is possible that novel situations may arise for which the automation was not designed. In these cases, the human operator's judgment may be more effective than the automation (Inagaki, 2000).

These examples do not necessarily mean that the operator should always control allocations. Scerbo (1996) cites several potential problems with operator-initiated invocations:

- Automation may need to change at a precise time at which the operator may be too busy to make the change.
- If there is a sudden increase in workload, the human operator may not be able to manage the automation correctly.
- As the human operator becomes fatigued, he or she will be less likely to take advantage of automation.
- In extreme cases of risk to the operator, the system might need to have authority if the operator is unable to maintain control (e.g., the GCAS described in the previous section).

However, in many cases, even if a computer maintains control, the human operator and system observers may still perceive the operator as having the ultimate decision responsibility. Consequently, from the operator's standpoint, he or she should manage automation invocations.

The problems that could result from assigning invocation authority to the computer or the human can vary in relevance depending on the context of the system being designed. For example, in a manufacturing system the causes and consequences of automation failures may be quite different from those occurring in an air traffic control (ATC) system. Both of these situations differ not only in the costs of failure, but also in who is perceived as being responsible (the automation designer or the human operator). Understanding how system context can affect the philosophy for determining invocation authority can help a designer manage these potential problems. Golikov and Kostin (1997) describe three approaches to FA based on which party bears the ultimate responsibility for the success of the system. These include a *human-centered* approach, a *machine-centered* approach and an *equivalent* approach to automation. Each of these approaches offers a different recommendation for the degree of human operator control of function allocations. Machine-centered approaches attempt to maximize the extent of system control through automation to the point where the human can be eliminated from the loop. In these systems, the designers rather than operators bear the ultimate responsibility for the system's success or failure. Machine-centered systems are often implemented in manufacturing facilities (Golikov & Kostin, 1997).

In the human-centered approach, the automated system offers assistance and serves as a tool to the human operators, who are ultimately in control of the system. These systems offer "semiautomatic control" instead of the fully automated scenario promoted by the machine-centered approach. The human operator is in charge of supervising and controlling the system, which requires the automation to be adaptable to human cognitive workload. Continuous participation with the system is required to manage novel or emergency situations when automation fails. This approach is common in modern aviation systems.

Golikov and Kostin (1997) do not believe either of these approaches is an adequate solution for modern systems. Human operators may be better suited for qualitative analysis in the case of novel situations than a fully automated system. However, there are some unexpected emergency situations in which a human operator may not be prepared to respond. In reaction to the potential shortcomings of these two approaches, the authors propose an *equivalent* approach, in which the designers and operators each share equal responsibility and work together to ensure the reliability and success of the system. The authors believe that modern integrated systems with complex controls require this equivalent approach, which would allow for resolution of both types of extreme situations. In the case of unexpected automation failures or novel situations, the human operator is best left in control, and in severe cases when the human operator cannot react adequately, he or she must be released from control by automation. In the latter case, the operator may regain control from the system when he or she is ready. Application of the equivalent approach can combine the advantages of the human and machine-centered approaches for managing fluctuating situations such as these, which occur in dynamic environments. This research suggests that there will be situations in modern systems, such as those found in aviation, when either the operator or the computer will need to be in control. It is therefore important for designers to understand the impacts of both authority conditions.

Hilburn, Molloy, Wong, and Parasuraman (1993) also considered exceptional situations when designing for adaptive automation. They conducted a series of experiments that examined operator performance when manipulating various aspects of AA, including authority over invocations. Citing Barnes and Grossman (1985), they used the terms *executive* and *emergency logic* to describe two perspectives regarding human and system-centered authority. When a system implements executive logic, the human operator has the final authority, while emergency logic places the control with the computer. At the time of their study, the authors noted a prevailing assertion that operator control was superior to system-initiated allocations. They performed an experiment in order to verify this assertion while observing the effects of executive and emergency logic within an adaptive system. Their experiment used the Multi-Attribute Task (MAT) battery, a PC-based simulation of flight tasks including simultaneous presentations of compensatory tracking, system monitoring, and fuel management tasks. Performance in the MAT is based on composite scores in each task, including monitoring reaction time and accuracy and errors in the tracking and fuel management tasks. In Hilburn et al. experiment, the fuel management and monitoring tasks were performed under manual control. Tracking was either automated or manual, depending on the experimental condition. Each participant completed four 20-minute sessions. In two of the sessions, tracking automation authority was under computer control, while in the other two sessions, the participant had control of automation shifts. Once automation was invoked (by the human or computer), the tracking task was performed by completely reliable automation for two minutes, then control of the tracking task was returned to manual. Using a model-based approach to AA, which uses a predefined scheme to determine the timing of DFAs, subjects shifted between manual and automated control at specific intervals and durations during four 20-minute sessions, either by computer or their own initiation. Hilburn et al. results confirmed a performance cost of automation shifts calculated across the three tasks (two monitoring and one tracking). However, these costs were reduced if the shifts were under operator control.

Expanding on these results, the authors describe situations requiring executive or emergency logic as two extremes of a continuum representing the control of a complex system. Between these two extremes, systems could provide recommendations for automated support, but grant final authority to the human operator, or the operator could request automated assistance from the system when he or she believes it is necessary. There are numerous other examples of intermediate stages along this continuum at which the human operator initiates, agrees to, or is directed by, automation. Rather than having a system that utilizes either executive or emergency logic, allocations could be the result of an operator agreement to a system suggestion, or vice versa (Hilburn, Molloy, Wong, and Parasuraman, 1993).

In a follow-on experiment, Hilburn et al. (1993) compared operator control under three different model- and performance-based approaches to AA system control. Using the MAT battery, three groups of subjects were exposed to shifts from automated to manual control in operation of the system based on different criteria. Only the monitoring task could be automated, while the tracking and fuel management tasks were always under manual control. One group of subjects experienced a model-based approach with emergency logic applied to invocation authority and system control was automatically switched to manual control in the fifth minute of a nine minute trial. The second and third groups experienced two types of performance-based approaches to adaptive automation. Subjects in the second group were forced into manual control if their performance fell below 57%. This represented emergency logic. The third group, which utilized executive logic, would also revert to manual control if their performance fell below this level, but could override the allocation if desired and remain under automated control. The results of this experiment indicated that automation benefits continued after participants returned to manual control. The type of adaptive logic (model- or performance-based and emergency or executive logic) did not reveal any clear advantages for any of the three conditions, although the authors acknowledge that additional analysis of the results could reveal a different conclusion. However, the general results of both studies by Hilburn et al. indicate that the locus of control in an AA system can have an effect on overall system performance.

In another study, Kaber and Riley (1999) also addressed the issue of human or automation initiated invocation of adaptive automation. This research assessed the usefulness of a secondary task measure for assessing workload and directing DFAs in a primary cognitive task. In their experiment, subjects performed two tasks: a primary dynamic control task, and a secondary gauge monitoring task. The primary task consisted of a "radar scope" display that presented randomly appearing square targets traveling at different speeds and directions toward the center of the display. The subjects were required to select and eliminate targets before they reached the center. The interface provided two levels of automation, including "manual control" and "blended decision-making." In the latter condition, the computer would first produce a list of decision options from which it would select a course of action. Implementation of the action required the consent of the human operator. The operator could either agree to the computer's selection, or select a course of action independently. The computer would then carry out the action selected by the operator (Kaber and Riley, 1999). The secondary task, which required subjects to maintain a moving pointer within a defined range on a fixed scale display, served as an objective measure of workload. Subjects were divided into two groups according to who was in charge of dynamic control allocations (the human or computer). The first group received mandates from the computer on when to invoke control allocations between the manual and automated control modes. The second group received control allocation suggestions from the computer, but was not required to make the allocations. The results showed that under manual control conditions, mandated subjects performed significantly better than the group receiving suggestions. Furthermore, subjects who received suggestions rather than mandates were more likely to accept automation, but less likely to agree to manual control. These results suggest that the additional cognitive requirements required for evaluating the decision to make an allocation may redirect the resources required for attending to the primary task.

These examples from the literature suggest that intermediate approaches to assigning authority in AA systems provide alternatives to static human- and machine-centered approaches to FA and the reliance on extreme executive or emergency logic to govern dynamic allocations. By modifying the invocation logic, a designer can manage the impact of DFAs on the operator. It is important to fully understand these impacts prior to implementing an AA system.

Although this review of literature points to a need for exploration of 'intermediate' levels of authority in invoking automation (such as suggestion and approval), there has only been limited empirical analysis in this area. Hilburn et al.

(1993) investigated the effects of human versus computer authority, but not at possible intermediate stages, such as approvals and suggestions for automation. Kaber and Riley's (1999) investigation focused on intermediate and high levels of authority. They compared the effects of computer mandates and computer suggestions on adaptive system performance. However, the scope of these studies has only considered the application of AA to a limited number of possible complex system functions (aspects of human-machine system information processing). It is possible that the levels of authority investigated by Kaber and Riley may have different effects on overall human-machine system performance when applied to various automated functions, such as detection tasks as compared to tasks requiring higher cognitive functions, like decision making. Intermediate stages of authority should be investigated fully in a generalizable context in order to understand the effects of authority for implementation in a variety of systems. Models for generalizing automation functions are discussed in the following section.

1.3 Automation of Stages of Human Information Processing

Wickens and Hollands (2000) state that automation can be designed in terms of how it assists the human operator in three different stages of information processing. In the first stage, which includes information acquisition and analysis, the system acquires, interprets, and/or integrates information, which replaces early human sensory processes, such as attention or perception (Wickens & Hollands, 2000). Examples include graphic displays such as electronic maps and ecological displays. Information presented by the automation at this stage supports decision making, the next of Wickens and Hollands' information processing stages. When decision making functions are automated, the automation offers or imposes constraints on choices available to the operator. An example of this type of automation would be an airborne traffic warning systems that advises pilots of a particular maneuver based on an emergency situation. In the final stage, execution, the automation carries out an action on behalf of the human operator. High levels of automation of any of these stages can be beneficial if its functions meet the expectations of the designer and the human operator (Wickens and Hollands, 2000).

Endsley and Kaber (1999) developed a similar stage model of human-machine system information processing by analyzing the common features of dynamic cognitive and psychomotor tasks common to the domains of ATC, piloting, advanced manufacturing, and teleoperations. From the common features, they identified four generic functions, including:

- Monitoring - scanning displays to perceive status
- Generating options - formulating options or strategies for achieving goals
- Selecting - deciding on an option or strategy
- Implementing - carrying out the chosen action

From these four functions a detailed taxonomy of ten levels of automaton was developed that divided the responsibilities for the functions between the human operator and the computer.

Endsley and Kaber (1999) applied their taxonomy to the performance of a dynamic control task, which yielded several notable results. First, subjects had better performance under levels of automation that combined human option generation with computer implementation as compared with automated option generation or fully manual control. Overall, human operators benefited most from physical implementation assistance and were somewhat hindered when assistance was provided with higher-level cognitive functions. Finally, generation of options resulted in worse performance when it was divided between the human and the computer than when it was performed by either one individually. These results show that various levels of automation, based on a stage model of human-machine system information processing, applied to a dynamic control task can have different effects on overall system performance.

In another study, Parasuraman, Sheridan, and Wickens (2000) introduced a model of *types* and *levels* of automation that also identified stages of information processing. In this model, type refers to the stage of information processing that is automated (information acquisition, information analysis, decision making, action implementation), while level indicates the degree of automation of a specific function. Automation can be designed according to a continuum ranging from completely manual control (low-level automation), in which all aspects of a function are under human control, to full automation (high-level automation), in which the machine controls the function and only the products of its operations are available to the human operator. Between these two extremes, types of automation can occur at varying levels to create innumerable combinations depending on task and system requirements.

For each type, Parasuraman et al. (2000) described the scope and examples of levels of automation. Automated information acquisition affects the sensing and registration of input data, which supports human sensory processes. Low levels of this type of automation might include data gathering, while higher levels could range from organizing the data to filtering the data so only a portion of it reaches the operator. Automating information analysis includes supporting higher cognitive functions such as working memory and inference. Low levels of this type of automation might involve providing task processing predictions to operators, while high levels could involve data integration (i.e., combining several data variables into a single value for decision processes). Automated decision and action selection includes generating hypotheses and selecting from among several alternatives. Sheridan's (1992) hierarchy of ten different levels of automation is relevant to this stage. The final automation type, action implementation, executes an action choice. In essence, it replaces the motor operation or command of the human operator (Parasuraman, Sheridan, and Wickens, 2000). An example of this type of automation is the motor commands required to correct an aircraft's altitude in the GCAS described in an earlier section.

The first two steps in Parasuraman et al. (2000) model for designing an automated system are to (1) recognize that automation varies by type and (2) decide which stage should be automated. When resolving the second step, the four automation types can be established at varying levels and in varying combinations for the design of an automated system. All or none of the human information processing stages can be automated at low, moderate, or high levels, depending on the contextual requirements of the system. Given the numerous available alternatives when designing an automated system, it is important to understand the effects of automation on the four general stages of human-machine system information processing in order to create technology that optimizes human performance.

The categories of automation described in this section are all based on generic models of human information processing. The four functions included in Parasuraman et al. (2000) model can be automated to augment or replace the different aspects of human information processing in complex system operation. However, they do not define any specific levels of automation, instead they only describe “low” and “high” levels. This makes it convenient to describe existing systems in terms of the model, but difficult to systematically examine the effects of these levels prior to system design.

The types of automation included in Parasuraman et al. (2000) model are similar to the stages described by Wickens and Hollands (except that information acquisition and analysis has been divided into two distinct processes) and nearly identical to the stages defined by Endsley and Kaber (1999). Endsley and Kaber, unlike Parasuraman et al., produced ten discrete levels of automation by combining types of automation and invocation authority to be used as a basis for systematically assessing the effects of these levels of automation on operator performance (in terms of workload and SA) and to make general recommendations for system design. However, none of these combinations isolated any of the four stages of information processing, so the advantages and disadvantages of each stage and type of authority is confounded within each level of automation.

The human factors literature includes several models of human-machine system automation defined on the basis of information processing models (Wickens and Hollands, 2000; Endsley and Kaber, 1999; and Parasuraman et al., 2000). Each of the complex functions identified by these models can be applied to systems in an adaptive manner. However, automating these functions may have varying effects in terms of operator performance, SA and workload. As described earlier, these effects may be further influenced by the interaction of various types of invocation authority that are used in AA systems. It is possible that different types of invocation authority may degrade or improve the effectiveness of AA as applied to, for example, information acquisition as compared to information analysis, etc.

At this point in time additional research is needed to further examine the performance effects of automating various stages of human-machine system information processing and different forms of authority over the invocation of automation. Systematically isolating automation of each stage of information processing (information acquisition, information analysis, decision making, and action implementation) and using different levels of invocation authority, as described in the previous section (computer mandating or computer suggesting) in a laboratory experiment may provide additional insight into the effectiveness of AA for complex systems. System designers could potentially use the resulting information to better understand in advance the effects of general types of automation and who is in charge during normal system operations.

1.4 Controlling Allocations

In addition to determining whether the locus of control should be given to the human or the computer and which system functions to automate, another important issue that must be addressed in designing AA systems is to determine the circumstances that trigger an allocation. An effective trigger should relieve the operator under high workload conditions and return control when operator workload returns to normal levels; thereby, maximizing the benefits of adaptive automation. Several methods for triggering dynamic control allocations have been proposed, including model-based approaches, physiological assessment, and performance-based approaches. The type or method of triggering DFAs is particularly important to this research because it is related to the authority over automation invocations. For example, performance-based systems typically operate under a computer authority that assesses operator functional state and makes decisions about types and levels of automation.

Triggers that incorporate models are based on expectations that workload will increase or decrease at a particular time or under predefined conditions. Hilburn et al. (1993) incorporated a model-based approach in order to determine the costs of frequent cyclings between periods of automated and manual control. Physiological measures can be used to infer operator mental workload while a task is being performed. In one study, Prinzel, Scerbo, Freeman, and Mikulka (1995) used arousal measurements recorded within an electro-encephalogram to trigger AA in a dynamic system. Their closed-loop system was able to successfully evaluate task workload and to predict appropriate system function allocations. Performance-based approaches to AA rely on observations of previous performance in a task to determine when periods of automation may be effectively initiated (Hilburn, Molloy, Wong, and Parasuraman, 1993). Another performance-based approach incorporates real-time assessments of workload using a secondary task measure. In these systems, performance variations in the secondary task serve as an indicator for levels of operator workload. The experiment performed by Kaber and Riley (1999) that addressed the issue of invocation authority of AA evaluated the efficacy of a secondary task measure as a triggering mechanism for adaptive automation. It was determined that basing primary task performance control allocations on secondary task performance can serve as an effective means for managing primary task workload.

In a recent study, Kaber, Wright, and Clamann (2002) utilized the same secondary task workload measure used in Kaber and Riley (1999) to compare the effects on human-machine performance of applying AA exclusively to four stages of information processing (information acquisition, information analysis, decision making, and action implementation) in a primary dynamic control task. The goal of this study was to examine the impact of AA on cognitive task performance and to determine which stages of human information processing are most conducive to adaptive automation. The dual-task scenario in this study involved subject performance of the primary dynamic control task and secondary gauge monitoring, simultaneously. Performance in the gauge monitoring task determined allocation of information processing functions to the human or the computer. The results provided evidence that the stage of human information processing that is automated has an impact on overall system performance. The results also further demonstrated the efficacy of a secondary task measure when integrated with an AA system for managing operator workload by effectively triggering dynamic function allocations.

The type of trigger used in an AA system directly affects the applicability of the authority condition. Regardless of type, each trigger provides some information to the automation that is designed to maximize system performance. Model-, physiological- and performance-based approaches for determining the timing of DFAs all imply some level of computer authority. The information developed based on the triggers used in these approaches is not necessarily known by the human operator and, therefore, must be presented either in the form of computerized advice or decision aiding, which constitutes suggestions, or in the form of a control mandate. The system works differently when the human is in control. Without computer aiding, the operator responds to his own suggestions, which can be influenced by other factors, such as heuristics, experience and training. Existing AA systems will therefore typically have some degree of computer authority.

1.5 Summary

Wickens and Hollands (2000) list three primary areas of concern for designers considering the implementation of adaptive automation. These areas include decisions on who controls the allocations, what functions are adapted, and how the system should infer the need for adaptation. These three areas have been addressed in previous sections and can be summarized as follows:

- Who controls the allocations:

  Designers who wish to achieve a locus of control that distributes responsibility to both the operator and the designer, as described by Golikov and Kostin (1997), must recognize where the machine's performance can surpass the operator's and where the operator can outperform the machine. These decisions can include intermediate stages of authority such as suggestion and agreement. They must also understand the performance effects of these authority conditions on the human who is left to operate or monitor the system.

- What functions are adapted:

  The literature identifies four stages of human information processing that can be used to define types of automation for complex system functions. Each stage of information processing can be automated or left under manual control. These stages include automation of information acquisition, information analysis, decision making, or action implementation functions. Each stage has its own unique effects on operator performance in terms of workload.

- How to infer the need for adaptation:

  A secondary task can be used to track and provide feedback on excess resources not allocated to the primary task. When resources are no longer needed by the primary task (due to a reduction in task workload, for example), they can be allocated to the secondary task, which can result in increases in secondary task performance. These fluctuations in secondary task performance can be used to control AA in the primary task as well as draw conclusions on aspects of the primary task that lead to the increases in workload. According to Kaber, et. al (2002), a secondary task performance measure can be used to control allocations in an AA system that automates one or more stages of information processing. Based on the previous discussion, this type of trigger is most compatible with a computer authority condition.

In a performance-based AA system, there are several interactions among these areas of concern for system designers, each potentially having an influence on the other as well as on operator workload. For example, there is the cyclical relationship of a closed-loop system (e. g., Prinzel, Scerbo, Freeman, & Mikulka, 1995; Kaber & Riley, 1999). When the system is in use, increases or decreases in operator workload may cause a secondary-task trigger to indicate that a function allocation is necessary. The trigger notifies the controlling authority (computer) of the task status, who then invokes the function allocation (automating a function or returning it to manual control). The system loop is then closed by this allocation affecting the operator workload (i.e., feedback on the cost or benefit of the allocation).

There may be other interactions within AA systems as well. There are the individual (main) effects of various authority conditions (Hilburn, Molloy, Wong and Parasuraman, 1993; Kaber and Riley, 1999) and of automating different stages of information processing (Kaber, Wright and Clamann, 2002), which may also produce an interaction effect on overall system performance. This interaction may dictate the combinations of authority and automation of human-machine system information processing that optimize operator performance. It is this interaction that is of primary interest for this study.

2 Problem Statement

When designing an adaptive system, it is important to understand the impact of the automation on the human operator. Adaptive automation using intermediate levels of automation has been shown to benefit human operators in terms of workload and SA and to result in performance that exceeds static automation or completely manual control (Harris, Hancock and Arthur, 1993). However, a review of the literature on AA reveals that there are several design dimensions/parameters, such as who controls DFAs and which functions are automated, on which there is limited knowledge and that could have a major effect on operator performance.

There are prevailing assumptions that some complex system operating conditions are more compatible with either operator or computer authority over function allocations, such as emergencies to which a human might not be able to respond, or novel situations not anticipated by the designers of expert systems (Scerbo, 1996). The literature shows that authority, or locus of control, in and of itself, may also have performance implications. Hilburn et al. (1993) found that operator control of invocations reduced the performance cost of model-based allocations in a tracking task more than computer-controlled allocations. However, the results of Hilburn et al. experiment comparing the performance effects of emergency and executive logic were inconclusive and may warrant additional investigation. Kaber and Riley's (1999) study with a performance-based system showed that computer-mandated control allocations caused subjects to perform better than when they simply received suggestions.

The four stages of human information processing identified in models of automation reviewed here can be used to classify the types of functions that can be automated in an adaptive system (Parasuraman, Wickens and Sheridan, 2000; Kaber and Endsley, in press) and there is evidence that humans may be more or less adaptable to flexible automation of different stages (Kaber, Wright, and Clamann, 2002). However, the interaction of the stage to which AA is applied and the form of invocation authority has not been investigated systematically. Kaber and Endsley (in press) investigated various levels of automation defined by combining automation of specific stages of human-machine system information processing with different forms of authority. This research provided information on the effects of combinations of automation types, but the interaction effects of the forms of authority and automation conditions were not isolated.

Although there has been some research on the effects of authority in performance-based AA systems including, for example, automated decision making (Kaber and Riley, 1999), there has not been a systematic investigation of the combination of the four stages of information processing identified in the general models of automation and intermediate levels of invocation authority. Given that authority and type of automation can each have an effect on performance, it is important to know if there is any systematic interaction between these two areas of concern for automated system designers. In a performance-based system, operator performance is measured and reported by the system; therefore, it is necessary for there to be some degree of computer authority.

The goal of this study was to assess the performance and workload effects of applying AA (using a performance-based approach) to the four stages of information processing described in Parasuraman et al. (2000) model and facilitating DFAs through two levels of computer authority (suggestion and mandate). Performing a systematic study of different levels of authority combined with the four types of AA using a performance-based approach was expected to demonstrate the existence of any interaction between these aspects of AA design. The results should provide additional insight for effective AA design in complex systems.

In specific, it was expected that higher levels of automation, such as information analysis and decision making would be more compatible with mandated allocations, while lower levels, such as information acquisition and action implementation, would be more effective under partial human control. The additional load imposed by the requirement of operator acceptance of DFAs was expected to adversely affect the more highly cognitive tasks. Conversely, the computer suggestion for the lower levels of automation could have prevented the operator from being "surprised" by the automation; thereby, reducing the need to readjust to the automated condition after an allocation.

3 Methods

3.1 Task

The experiment included two computer-based tasks, a dynamic control task (Multitask©) and a secondary gauge-monitoring task. Both of these tasks were modified versions of tasks employed by Endsley and Kaber (1999), Kaber and Riley (1999), and Kaber et al. (2002) in studies of the performance and workload effects of AA in dynamic control tasks and the effectiveness of a psychophysical-based approach to AA under different forms of authority for managing operator workload.

3.1.1 Multitask

Multitask is a dynamic control simulation analogous to a low-fidelity ATC task. Multitask was originally designed to incorporate common features of dynamic control tasks such as piloting and ATC (i.e., multiple competing goals, multiple tasks of varying relevance competing for the operator's attention, and high demands under limited time resources). It has been used successfully in several recent studies (e.g., Bolstad and Endsley, 2000; Endsley and Kaber, 1999; Kaber and Riley, 1999) to investigate, for example, the effects of intermediate levels of automation and the usefulness of a secondary task measure for assessing workload and directing DFAs. A high-level task analysis of Multitask appears in Appendix A.

The Multitask interface (see Figure 3.1) presented several targets simultaneously on a radarscope display. Targets were represented by icons of three different types of aircraft (military, commercial, and private; see Figure 3.2) traveling at different speeds toward the center of the display. The speed of each target depended on its type. Military aircraft had the highest maximum speed, followed by commercial aircraft and then private aircraft. There was some variation of speed within each type of aircraft. Targets required between approximately 60 and 120 seconds to reach the center of the display subsequent to their appearance on the screen.

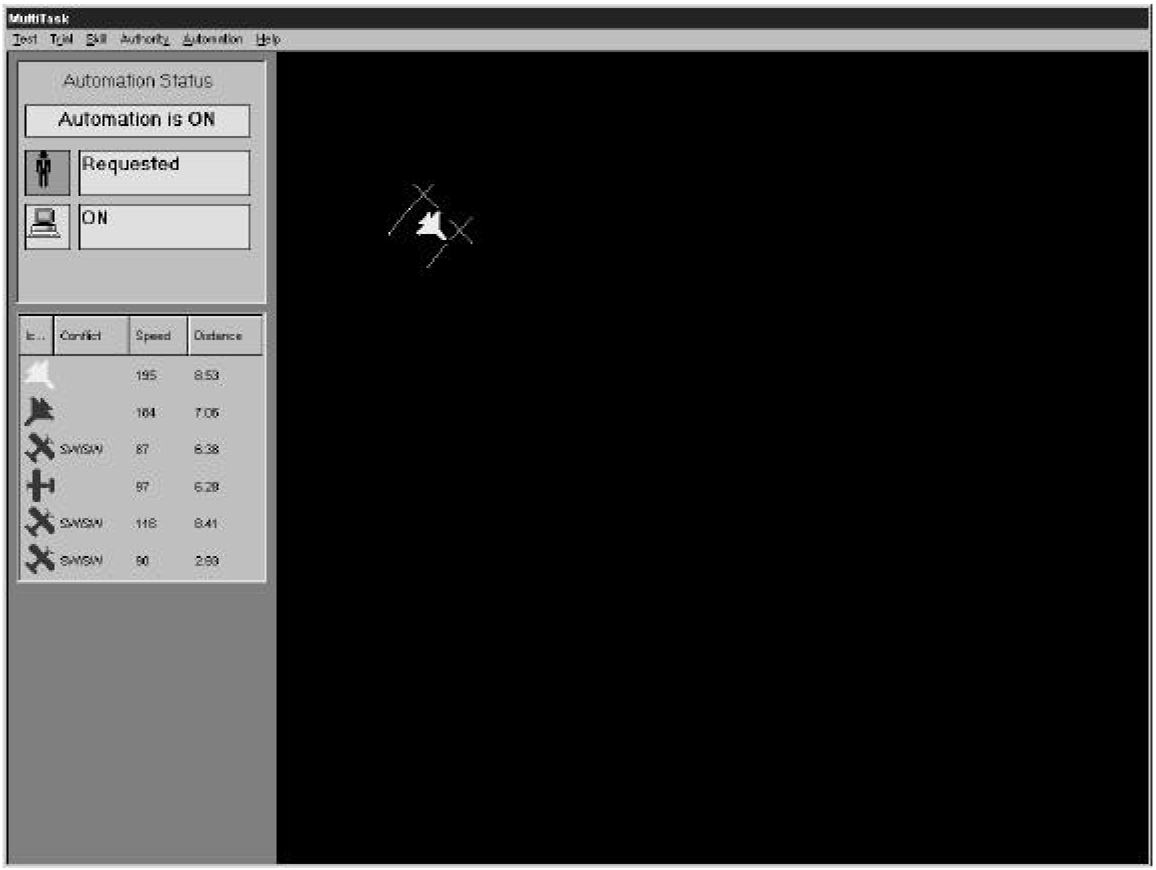
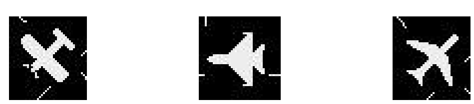


Figure 3. 1 Multitask with automated information analysis.



Private Military Commercial

Figure 3. 2 Various aircraft types.

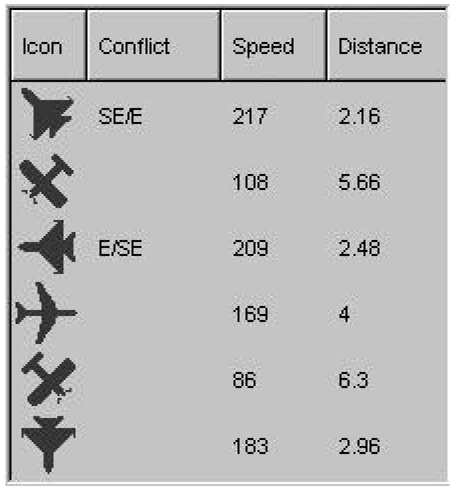
The operator's task was to locate and "clear" aircraft for landing before they reached the center of the display (airport) or collided with another aircraft. Clearing an aircraft required two steps: (1) establishing a communication link and (2) issuing a clearance. To establish a communication link, participants moved a cursor to an aircraft using a mouse controller, and then pressed the left mouse button. The aircraft flashed for several seconds, signifying a processing stage. After the aircraft stopped flashing, the subject clicked the aircraft with the right mouse button. The aircraft flashed again, this time for a significantly longer period. Once the aircraft stopped flashing the second time, a tone was played indicating that the clearance had been issued and the aircraft could safely approach the airport. Depending on the stage of processing, the color of aircraft may have been red (a new target), flashing red (a communication link pending), yellow (a communication link established), flashing yellow (a clearance pending), or green (a clearance issued). Clearing each aircraft required at least 30 seconds (seven seconds for establishing a link followed by 23 seconds for the clearance) and the operator could clear multiple aircraft simultaneously.

During all training and experiment trials, the majority of the Multitask radar display was not visible to the participant (as shown in the central portion of Figure 3.1). Only a small area was made visible through a portal, or keyhole, that could be moved by the participant in horizontal, vertical, and diagonal directions using the numeric keypad as part of the keyboard. In order to establish communication links or clear aircraft, participants had to first locate them using the portal. Audio feedback was provided when aircraft collided, reached the center of the display without being issued a clearance, or when a clearance was issued.

The version of Multitask used for this study provided one-of-five different modes of automated assistance to a user. Each mode was designed to automate a single stage of information processing and is described as follows:

1. Manual - No additional assistance was provided.
2. Information Acquisition - The computer controlled the movement of the portal, which moved in a circular motion toward the center of the display. Each full cycle of the portal covered the entire display. By pressing a key, participants could optionally lock the portal on to aircraft as the portal's automatic movements exposed them. This feature is referred to as 'automatic tracking.' With the tracking feature enabled, the portal moved in its regular pattern around the display until it revealed any part of an aircraft that had not yet been selected for clearance. Once an aircraft was located, the portal centered itself on the aircraft and continued to follow the aircraft's path until the participant started to issue a clearance (i.e., the aircraft was flashing yellow) or released the portal from the aircraft by pressing the enter key. This type of automation is similar to military ATC, which also uses a fixed-scanning pattern radar and can also optionally lock on to specific targets (Parasuraman, Sheridan, and Wickens, 2000).
3. Information Analysis - A decision aid was presented on the display showing the layout of all aircraft, including type, direction, speed, distance from the center of the display, stage of processing, and a projection on whether or not the aircraft might have been involved in a collision (see Figure 3.3). Information on each aircraft was presented in a random order. Automation of this stage is analogous

to the converging runway display aid used in ATC that automatically projects the approach paths for multiple aircraft (Parasuraman, Sheridan, and Wickens, 2000).



Icon	Conflict	Speed	Distance
	SE/E	217	2.16
		108	5.66
	E/SE	209	2.48
		169	4
		86	6.3
		183	2.96

Figure 3. 3 Information analysis decision aid.

4. Decision Making - A decision aid similar to the one used in the information analysis condition was presented, but without the speed, collision, and distance information (see Figure 3.4). Instead, the decision aid sorted aircraft for processing according to priority, from the top to bottom of the display. Priority was calculated by the computer. The highest priority was given to aircraft on collision courses with other aircraft, then to those aircraft closest to the center of the screen. The Multitask algorithm that determined priority did not always provide the best recommendation. For example, since the recommendations were based, in part, on proximity to the center of the screen, it was possible for the computer to prioritize an aircraft that could not be cleared in the time remaining before it met the due date. Parasuraman and Riley (1997) describe proposed

decision aids that provide suggestions to air traffic controllers for resolving potential aircraft collisions as an example of automated decision making in ATC.

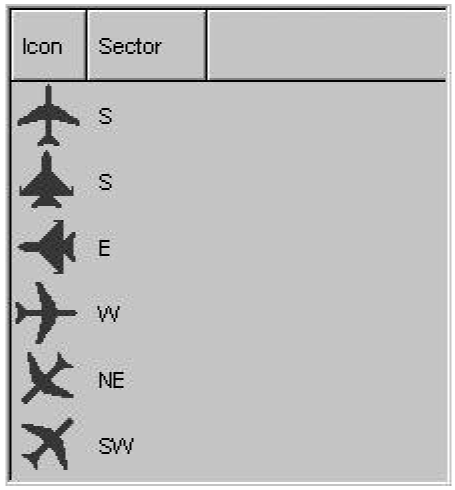


Figure 3. 4 Decision making decision aid.

5. Action Implementation – This form of automation presented a decision aid listing only the number of aircraft and their stages of processing (see Figure 3.5). In this mode, the clearance was issued automatically after the communication link was established. Participants were only required to click aircraft once to issue a clearance. This final stage of automation was similar to the automated transfer of an aircraft from one sector to another in commercial ATC systems that can be completed with a single press of a button once the controller has made the decision for the transfer (Parasuraman, Sheridan, and Wickens, 2000).

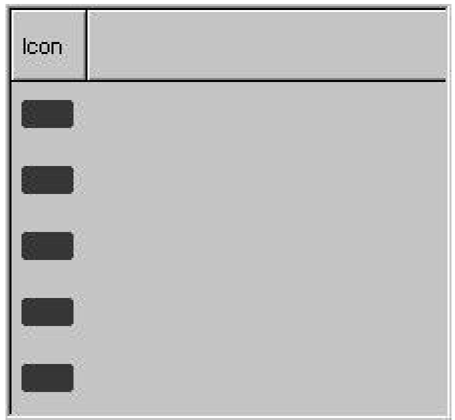


Figure 3. 5 Action implementation decision aid.

In addition to automation modes, there were also four types of automation invocation authority available in the Multitask simulation. These were determined by the locus of control (computer or human) and the level of authority assigned (mandate or suggest). When the authority level was "mandate," the computer or human (depending on the condition assigned) had full control of automation and manual control invocations. Computer allocations occurred at the precise time a "performance event" occurred, such as an operator's performance in the secondary task falling below a predetermined criterion, or a substantial increase in the level of operator workload due to the primary task. Optionally, the system could be configured to mandate allocations at timed intervals instead of according to operator performance. Human allocations of automation or manual control occurred when the operator pressed the space bar on a keyboard integrated with the system running the multitask simulation. When the authority level was "suggest," these same events (human presses the space bar or the computer detects a critical change in performance) represented either a suggestion or an approval of a control allocation, depending on who had authority. The various settings for locus of control and level of authority had the following implications for Multitask performance:

1. Computer Mandate – Computer had full control of allocations. The state of the automation was determined, based on either on operator performance events or a predetermined schedule of manual and automated control periods (e.g., automation cycles every two minutes). Human input had no effect on triggering automation.
2. Computer Suggest – Computer had to initiate allocations. Operator performance events changed the state of the automation to "automation recommended" or "return to manual recommended." Human had to agree to automation or manual control before an allocation could occur. Input by the human prior to a computer suggestion was ignored.
3. Human Mandate – Human had full control of allocations. Human input determined the state of the automation. Operator performance events had no effect on triggering automation.
4. Human Suggest – Human had to initiate allocations. Human input changed the state of the automation to "automation recommended" or "return to manual recommended." Computer had to agree to automation or manual control before an allocation could occur. The computer would only respond to an operator performance event if automation had been suggested by the human.

In order to make a direct comparison with Kaber and Riley's (1999) research, the present study considered the first two options - computer mandate and computer suggest. Feedback on the current state of the locus of control and automation suggestions was provided by an additional window on the Multitask display (see Figure 3.6 and upper-left corner of Figure 3.1). If the level of the authority was “suggest” (Settings 2 and 4 described above), icons of a computer and a person provided feedback on the state of the recommendation or agreement. The background color of the icon reflected the state of the automation request. “Blue” designated a request for manual control while “yellow” signified a request for automation. An allocation occurred when both colors matched (yellow for automation and blue for manual control).

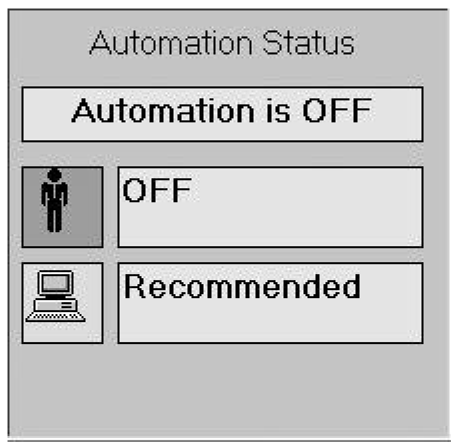


Figure 3. 6 Automation Status Indicator.

Under each condition, the participant’s goal was to maximize task reward points. There were no partial rewards assigned in the task. One point was awarded for every fully cleared aircraft and an additional point was added if a cleared aircraft was on a collision course with another aircraft; that is, two or more aircraft were within a predefined “buffer zone” of each other. This buffer zone was the airspace surrounding each aircraft designated as the safe distance from surrounding aircraft.

3.1.2 Gauge Monitor

The gauge-monitoring task presented a fixed-scale display with a randomly moving pointer (see Figure 3.7). The fixed-scale included a color-coded "acceptable" central region and both an upper and a lower "unacceptable" region. The acceptable region was "green" and the unacceptable regions were colored "red". Participants were required to monitor vertical pointer movements in order to detect when a deviation occurred into either of the unacceptable regions. Participants could correct for pointer deviations by pressing keys on the keyboard, which centered the pointer on the fixed scale.

Although the pointer generally moved in a random manner, it was designed to not stay in the unacceptable region for extended periods of time. This was done in order to maintain a consistent number of deviations among subjects. If the pointer remained in the unacceptable range for more than three seconds, it would move back into the acceptable range on its own and register a "miss." This design produced approximately six deviations per minute. Three seconds was determined in previous pilot testing to be ample for scanning between the primary task and the gauge.

Performance in the gauge task was recorded as a ratio of the number of hits to the number of unacceptable pointer deviations (i.e., the hit-to-signal ratio). Significant variation in this performance measure resulted in dynamic allocations of system information processing functions to the Multitask computer system or to the operator, as established by previous research (Kaber and Riley, 1999).

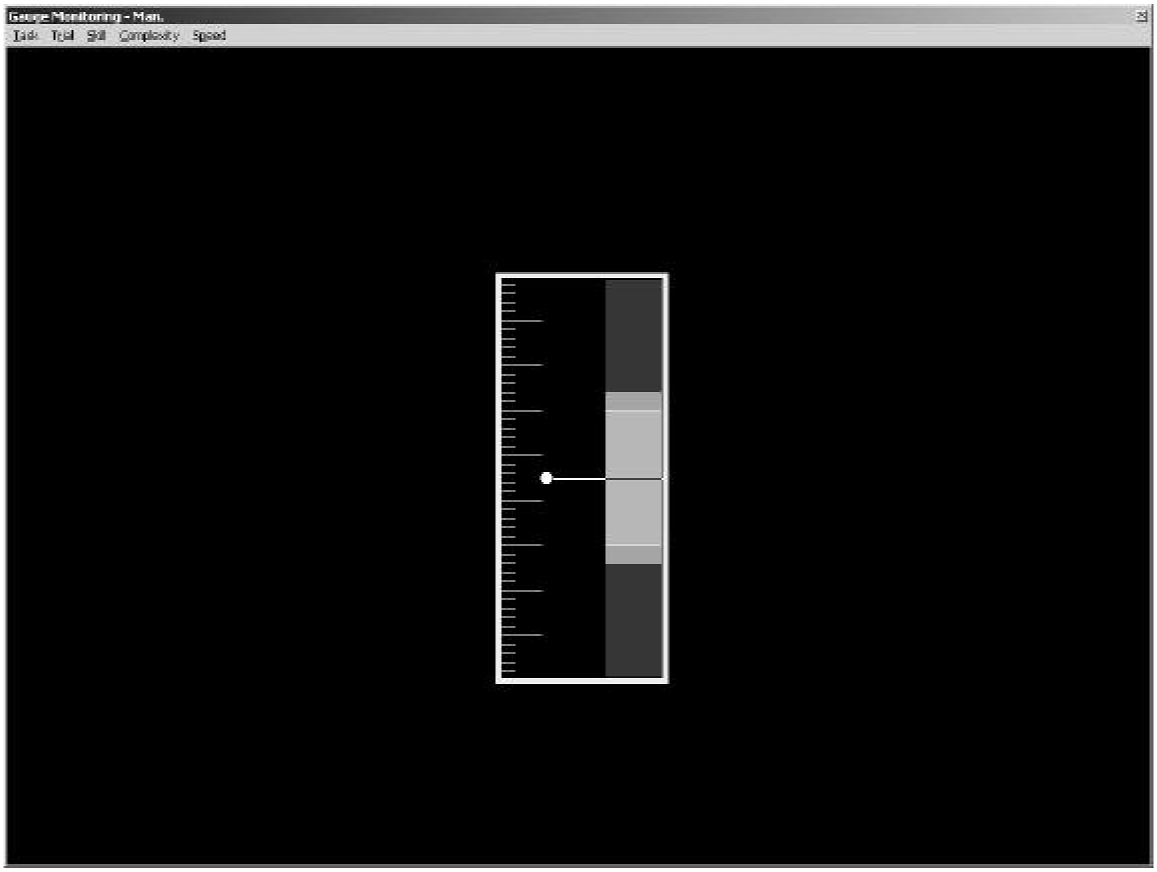


Figure 3. 7 Gauge Monitor.

Table 3.1 contains a summary of the controls used in Multitask and the Gauge Monitor.

Table 3. 1 Summary of Task Controls.

Control	Application	Purpose
Enter key	Multitask	Enable/disable automated tracking (available only when information acquisition is automated)
Left Mouse Button	Multitask	Establish communication link
Right Mouse Button	Multitask	Issue clearance
Numeric keys 1- 4, and 6 - 9	Multitask	Move portal (except automated information acquisition)
Space Bar	Multitask	Toggle automation
Shift key	Gauge	Correct for downward deviations
Control key	Gauge	Correct for upward deviations

3.2 Participants

Forty-eight subjects were recruited for participation in this experiment from the North Carolina State University student population. They ranged in age from 19 to 34 years. All were required to have 20/20 or corrected to normal vision and some degree of windows-based computer or video game experience. All subjects were provided with $15 total compensation for their participation. There was an additional incentive in the form of a $30 gift certificate awarded to the participant who achieved the best overall performance during the experiment.

Of the 48 persons who volunteered for the study, only 40 subjects actually experienced AA (dynamic control allocations) during test trials based on their performance in the gauge monitoring task (objective workload). The task load posed in the Multitask simulation and the nature of the automation may have also been factors in whether subjects experienced AA. Since the specific objective of this research was to examine the effects of AA applied to various information processing functions, observations on five of the original 40 subjects recruited for the experiment, who did not experience AA, were not included in the overall experimental data set. Consequently, five additional subjects were recruited, three of whom also did not experience AA. The final three subjects that were recruited for the experiment all experienced AA, yielding the necessary number of observations on the performance measures for testing each experimental condition.

3.3 Equipment and Environment

The equipment consisted of a single computer workstation with 512 MB of RAM and a 2.0 GHz Pentium IV processor. Multitask and the gauge-monitoring task were presented across two 17-inch digital flat panel displays. A 104-key Dell Quiet-Key keyboard and a Microsoft IntelliMouse 2-D mouse controller were used to control of both tasks. Harman/Kardon speakers provided various audio cues.

The experiment took place in a single office room with the lights off in order to keep the participant focused on the tasks. Following the training sessions, it was expected that each participant would be familiar enough with the controls to perform the tasks in low-light conditions. In addition, there was sufficient ambient light from the two computer monitors to view the keyboard in case the participant needed to reorient him or herself during the experiment. The experimenter was seated out of view of the participant and took particular care not to distract any subjects during training or experimental trials and keep them focused on the tasks.

3.4 Experimental Design

This experiment followed a 4×2 mixed design including the four types of automation (information acquisition, information analysis, decision making, and action implementation) manipulated as a between-subjects variable and the two types of authority (computer mandate or suggest) controlled as a within subjects variable. A manual control condition was also investigated as part of the study in order to serve as a basis for comparison of the effectiveness of the various AA conditions. Eight participants were used for each of the five automation conditions and the manual control condition, for a total of 40 subjects. This was consistent with an earlier experiment that examined the effects of automation on different stages of information processing and demonstrated significant effects on performance and operator workload (Kaber, Wright and Clamann, 2002). Each participant experienced one type of automation under both types of authority in separate trials. The order of presentation of the two authority conditions varied from subject to subject. A summary of the distribution of the experimental conditions for each participant appears in the data collection table (see Table 3.2).

Table 3.2 Data Collection Table.

Type of automation										
Manual control			Information acquisition		Information analysis		Decision making		Action implementation	
Subjects										
Trial	1, 11, 21, 31	2, 12, 22, 32	3, 13, 23, 33	4, 14, 24, 34	5, 15, 25, 35	6, 16, 26, 36	7, 17, 27, 37	8, 18, 28, 38	9, 19, 29, 39	10, 20, 30, 40
1	N/A	N/A	Suggest	Mandate	Suggest	Mandate	Suggest	Mandate	Suggest	Mandate
2	N/A	N/A	Mandate	Suggest	Mandate	Suggest	Mandate	Suggest	Mandate	Suggest

The interaction of automation invocation authority and type of automation in the AA system was investigated through DFAs during the two experimental trials. Each participant began the two test trials under completely manual control and subsequent allocations of automation occurred according to a performance-based approach, as Kaber and Riley (1999) did. Significant changes in workload, as captured by the gauge monitor, resulted in computer mandates or suggestions for DFAs in Multitask. During periods of automation, the computer was responsible for one of the four stages of information processing, while the participant was still required to attend to the other three. During manual control periods, the participant did not receive any additional assistance clearing aircraft in Multitask.

3.5 Variables

The independent variables in this experiment included the form of automation, with the four conditions of information acquisition automation, information analysis automation, decision making automation, and action implementation automation (observations were also recorded on the manual control condition). The two authority conditions (computer mandate or suggestion) were also considered as a predictor variable. The dependent variables included the ratio of cleared aircraft to total aircraft, the ratio of aircraft collisions prevented to potential collisions reported (i.e., aircraft within the defined “buffer zone” of each other) and the hit-to-signal ratio recorded for the gauge monitoring task.

A subjective assessment of operator workload was performed using measures for mental and temporal demand. Subjects rated temporal demand (time pressure) using a visual analog scale (VAS) with anchors of "none" and "maximum." Subjects designated a rating by marking an "X" on the VAS at the position at which they felt best represented the level of time stress experienced in a trial. Mental workload levels were rated using an adaptation of the decision-tree as part of the Modified Cooper-Harper (MCH) scale. The MCH was developed for assessing mental workload in a dual-task scenario and has been validated as a reliable and robust measure of cognitive load by Wierwille and Casali (1983). Since the Multitask interface was purposefully designed to maintain high task workload levels, suggestions regarding interface design that are included in the MCH might have been considered leading and were not relevant to the study. The adapted scale was therefore modified from the original by removing several features that made suggestions regarding interface design (e.g., "Major deficiencies, interface redesign is strongly recommended"). The various forms for measuring mental and temporal demand appear in Appendices C and D, respectively.

3.6 Procedures

Each experimental session began with a brief introduction to the study, during which participants were given an overview of the procedures, including the goals and durations of the training and experimental periods as well as descriptions of the equipment. Participants then received specific instructions for Multitask and completed two 15-minute training sessions in the simulation, the first under completely manual control and the second with the assistance of one automated information processing function, as assigned to the subject in advance.

Following a five-minute break, participants received instruction on the gauge monitoring task and a five-minute training session. They were then required to complete a 24-minute dual-task training session including the gauge monitoring task and Multitask. Participants were instructed to distribute their attention evenly across both tasks, but to ensure that they addressed the targets presented on the Multitask display. During this session, automation of the information processing function cycled on and off every two minutes, beginning with manual control. The computer mandated all automation allocations and returns to manual control (In order to be consistent in the delivery of the experiment, subjects in the manual control condition experienced the same training protocol as subjects assigned to the various types of automation.).

During the dual-task training session, the average and the standard deviation (SD) of subject gauge-monitoring performance (hit-to-signal ratio on pointer deviations) were recorded during the latter 20 minutes of the training trial. These two figures were used as a basis for computer mandates or recommendations of shifts to automated or manual control. During the experimental trials, if average performance in the gauge task for a one-minute period fell below one SD of average task performance, the computer recommended or mandated automation of the information processing function. After switching to automated control, if the participant's performance increased to one SD above average or to a hit-to-signal ratio of 1.0 (perfect performance in the gauge task), the computer recommended a return to manual control or manual control was mandated. These criteria were based on those used by Kaber et al. (2002) for managing suggested and mandated allocations in a similar closed-loop system. Prior to beginning the experimental trials, participants received a 10-minute break.

Immediately prior to each experimental trial, participants were exposed to a two-minute demonstration of the authority condition to be used in the subsequent trial. During these sessions, the computer attempted to communicate to the participant a recommendation or mandate every 15 seconds. The participant was able to observe the effects of the various combinations of suggestions and mandates under both authority conditions. Both of the experimental trials lasted 20 minutes, during which automation allocations occurred based on human agreement to a computer suggestion or computer mandate, depending on the authority condition. Computer decisions to suggest or mandate automation were based on increases or decreases in gauge performance during the experimental trials relative to the performance baseline established in the dual-task training session.

Subjects assigned to the manual control condition also participated in two 20-minute trials; however, no automated control allocations occurred during these tests. *Table 3.3* provides a summary of the steps as part of the experiment that were performed with all participants. The complete set of instructions is also presented in Appendix B.

Table 3. 3 Experimental Procedures.

Step	**Duration**
Introduction and equipment familiarization	15 min
Multitask training under manual control	15 min
Multitask training including automation condition, followed by a five-minute break	20 min
Gauge monitor training	5 min
Dual-task (Multitask and gauge monitor) training with two-minute computer mandated automation cycles, followed by a ten-minute break	34 min
Multitask authority condition training	2 min
Experimental automation trial (or manual control condition)	20 min
Multitask authority condition training	2 min
Experimental automation trial (or manual control condition)	20 min
Debriefing	5 min
Total:	133 min

3.7 Hypothesis

It was expected that human-initiated invocations of DFAs would yield a smaller cost in terms of primary task performance than system-initiated invocations under all automation conditions. Computer mandates for automated and manual control did not provide feedback that might be important for effective interaction with the primary task (i.e., cues on control mode changes or timing information). Following a mandate, subjects were expected to have to reorient themselves to the system, while the suggestion condition allowed a subject to time his or her allocations strategically. Such a result would have been consistent with Hilburn et al. (1993), who found that operator initiated invocations imposed less of a performance cost than those initiated by the computer.

Furthermore, it was considered likely that operator effort and concentration required to respond to suggestions would impose an additional burden on performance of the higher cognitive functions (i.e., information analysis and decision making). In this authority condition, the participant had to decide whether or not to respond to suggestions, a task not required in the computer mandate condition. This could have resulted in an increase in primary task workload and, therefore, decreased secondary task performance. This would have been consistent with observations made by Kaber et al. (2002) that additional displays as part of an experimental condition in Multitask may require more cognitive processing.

4 Data Analysis

All statistical analyses were performed using SAS. They included multi-way analyses of variance (ANOVA) applied to the dependent variables to investigate gauge monitoring performance and primary task performance. The simple ANOVA model for the experiment can be written as follows:

$$Y_{ij} = \mu + TA_i + IA_j + \varepsilon$$

where

Y_{ij} = the response variable (e.g., workload)

TA_i = Type of automation

IA_j = Invocation authority

ε = Experimental error

The full model, including the interaction terms, can be written as:

$$Y_{ijk} = \mu + TA_i + IA_j + SUB(TA)_{k(i)} + TA*IA_{ij} + IA*SUB(TA)_{jk(i)}$$

where,

SUB(TA) = subject nested within TA

for,

i = 1, 2, 3, 4

j = 1, 2

k = 1, 2, 3, 4, 5, 6, 7, 8

An alpha level of 0.05 was used to identify any significant main effects of TA or IA and the presence of any significant interactions. Further investigation of significant predictors was conducted using post-hoc tests, specifically Tukey's Honestly Significant Difference (HSD) test with an alpha criterion of 0.05.

Correlation analyses were also conducted on the various response measures recorded during the experiment, including: (1) primary task performance measures and the objective measure of workload (secondary task performance); (2) objective and subjective measures of workload; and (3) subjective workload measures and primary task performance. Pearson Product-Moment coefficients were calculated to identify any positive or negative linear associations of the responses. The relation of primary and secondary task performance was examined in order to determine whether Multitask performance could be explained in terms of operator workload, as influenced by the computer assistance. Second, the degree of linear association between the subjective and objective measures of workload was determined in order to assess the accuracy of subject perceptions of task load and to indirectly validate the secondary task as an objective indicator of workload. Correlations were also calculated on the subjective assessments of workload (mental and temporal demand) and performance in the primary task in order to identify any relationships between actual performance and perceived workload across the various modes of automation. The SAS PROC CORR procedure was used to establish the statistical significance of the correlations of interest to the study.

The various response measures were organized in several different data sets in order to make comparisons of types of automation and authority and to reveal if the participants' performance during periods of manual control was influenced by automation (i.e., whether a carry-over effect occurred). Four different data sets were constructed for these analyses. The first set included observations on Multitask and gauge monitoring performance that were averaged for each subject across periods of automated assistance during each trial. This produced a single score for automation for each trial. The second data set included average performance for each participant across periods of manual control during each trial. This second set was further divided into two other data subsets based on the authority condition (mandate or suggest) and they included observations on the average performance of subjects assigned to the manual control condition. The data sets were used to isolate the effect of automation on manual control under specific types of authority and to make comparison with the control condition.

Based on this handling of the data, the set used to evaluate the effect of automation under both types of authority included 64 observations (4 automation types $\times$ 8 subjects $\times$ 2 trials), while those data sets used to evaluate the manual performance during AA under a specific authority condition (mandated or suggested) included only 40 observations (5 automation types (including manual control condition) $\times$ 8 subjects $\times$ 1 trial). In the latter data sets, the subject performance in the second trial as part of the manual control condition was considered.

In general, analyses were performed on 12 response measures, including: (1) targets cleared in the primary task while under automated control (64 observations); (2) targets cleared in the primary task while under manual control (64 observations); (3) conflicts resolved in the primary task while under automated control (64 observations); (4) conflicts resolved in the primary task while under manual control (64 observations); (5) targets cleared in the primary task while under automated control with mandated automation (40 observations (with reference to the subsets defined above)); (6) targets cleared in the primary task while under manual control with suggested automation (40 observations (with reference to the subsets defined above)); (7) conflicts resolved in the primary task while under automated control with mandated automation (40 observations (with reference to the subsets defined above)); (8) conflicts resolved in the primary task while under manual control with suggested automation (40 observations (with reference to the subsets defined above)); (9) gauge-monitoring performance while under manual control in the primary task (64 observations); (10) gauge-monitoring performance while under automated control in the primary task (64 observations); (11) gauge-monitoring performance while under manual control in the primary task with mandated automation (40 observations (with reference to the subsets defined above)); and (12) gauge-monitoring performance while under manual control in the primary task with suggested automation (40 observations (with reference to the subsets defined above)).

A Shapiro-Wilks test was conducted on each performance measure in order to determine whether each data set conformed to the normality assumption of the ANOVA. Beyond this, plots of residuals against the settings of the independent variables were used to assess the constant-variance assumption. In the event that the test produced a significant result ($p<=0.05$) or the plots revealed substantial instability in the variance of a response, data transformations on the response were considered in order to possibly achieve normality and to stabilize the variance. The transforms included square roots, logarithms, and reciprocals of the measures. If transforms were not successful in resolving ANOVA assumption violations or simply could not be logically applied (e.g., logarithm of a zero value), then a non-parametric method was applied to the data in lieu of the ANOVA. The Kruskal-Wallis test was identified as a comparable non-parametric test and was selected for these specific situations based on its effectiveness in analyzing non-normal data due to floor or ceiling effects, or due to outliers (Conover, 1980).

5 Results

5.1 Objective Performance Measures

5.1.1 Primary Task Performance

An ANOVA on Multitask performance during periods of automation based on the ratio of cleared targets to presented targets revealed significant effects due to both automation type ($F(3,48)=3.37$, $p=0.026$) and type of authority ($F(1,48)=7.29$, $p=0.010$). Individual differences were not significant in influencing this performance measure. A post-hoc analysis of these effects using Tukey's HSD indicated subjects performing the primary task with automated information acquisition performed significantly better ($p<0.05$) than those with automated decision making. The post-hoc analysis further indicated that those subjects receiving automation suggestions performed significantly better ($p<0.05$) than those for whom automation was mandated.

A Shapiro-Wilks test for normality on the resolved conflict data returned a significant result ($p<0.024$) and there appeared to be a substantial departure of the residuals (for the statistical model) from a linear trend in the normal probability plot. Consequently, various transforms were applied to the dependent measure in order to meet the assumptions of the ANOVA, including a square root function. Unfortunately, the transform of conflicts resolved was not successful and the Shapiro-Wilks test remained significant. As a result, nonparametric methods were used to assess the automation type and invocation authority affects on conflict resolutions. A Kruskal-Wallis test revealed significant differences due to automation type ($T(3)=11.30$, $p=0.010$). Additional applications of the Kruskal-Wallis to pairs of automation conditions indicated significant differences between information acquisition and action implementation ($T(1)=9.34$, p=0.002) and between information analysis and action implementation (T(1)=11.40, p=0.001). For both tests, subjects performed worse under automated action implementation in terms of conflicts resolved.

Figure 5.1 summarizes the mean performance based on cleared targets and resolved conflicts by automation type. Figure 5.2 presents the mean of cleared targets and resolved conflicts by invocation authority.

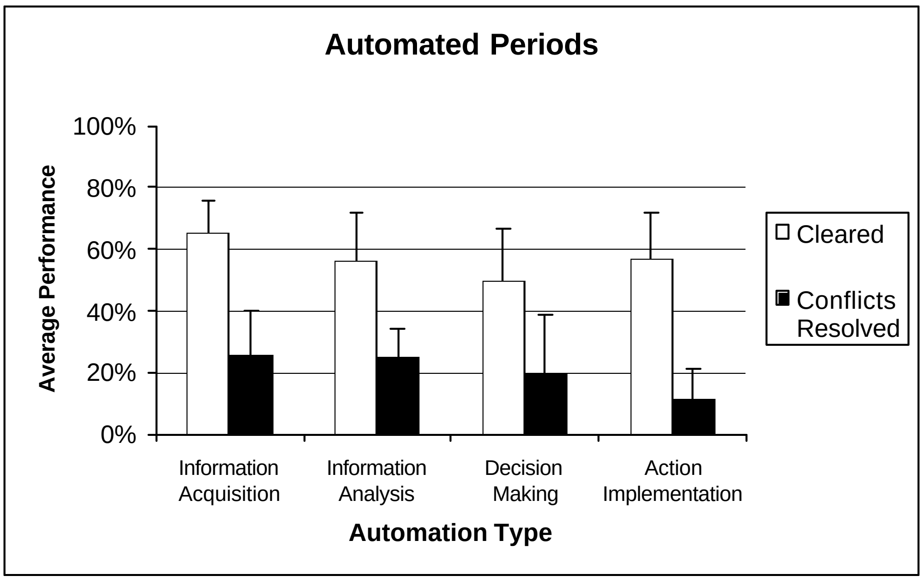


Figure 5. 1 Primary task performance during automation by automation type.

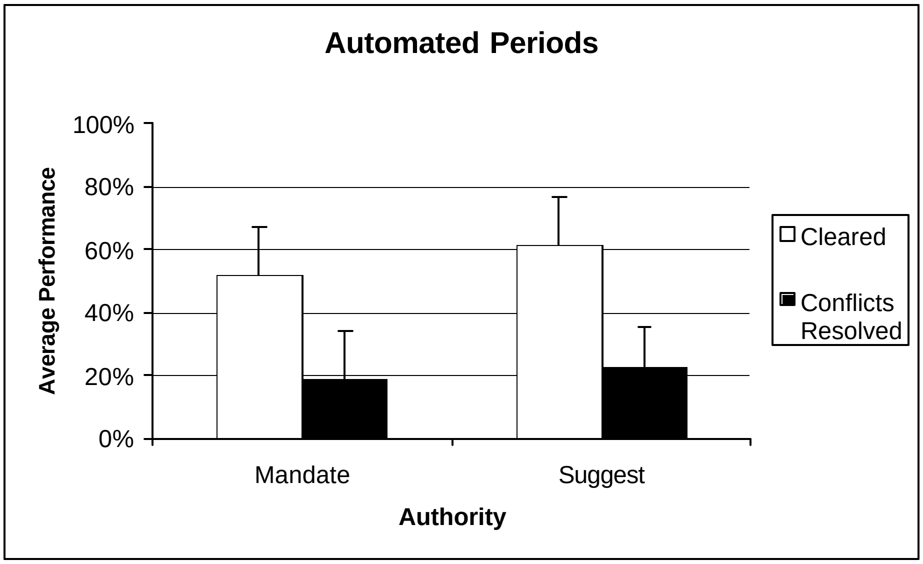


Figure 5. 2 Primary task performance during automation by authority.

An ANOVA on Multitask performance based on targets cleared during periods of manual control indicated a trend in the data due to the authority condition, which was on the edge of significance (F(1,48)=3.69, p=0.061). A post-hoc analysis of this trend with an alpha level of 0.10 suggested that subjects presented with automation mandates may have performed better in terms of targets cleared than those for whom the automation was suggested; however, these observations were not significant based on the criterion defined for the study.

A Shapiro-Wilks test for normality on the conflicts resolved during manual control was significant (p<0.001) and a normal probability plot revealed a non-linear trend. Consequently, transforms were applied to the response measure but, as with the data collected during the automation periods, they were not successful in correcting for the ANOVA assumption violations. Therefore, a nonparametric analysis was conducted using the Kruskal-Wallis, which revealed a significant effect due to the authority condition (T(1)=13.048, p<0.001). Specifically, subject performance under automation mandates was significantly superior ($p<0.001$) to suggestions. Figure 5.3 and 5.4 summarize the mean primary task performance measures by automation type and invocation authority, respectively.

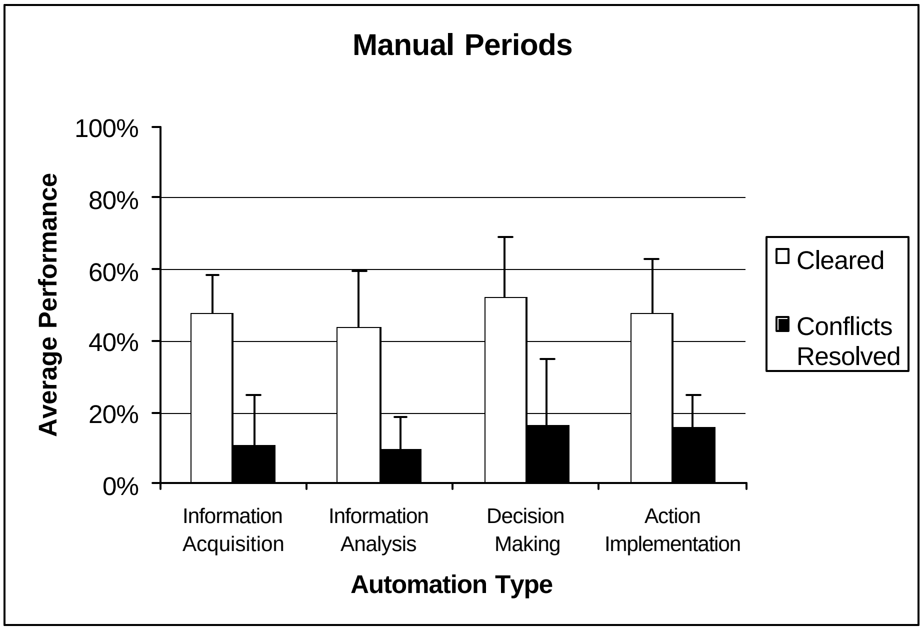


Figure 5. 3 Primary task performance during manual control by automation type.

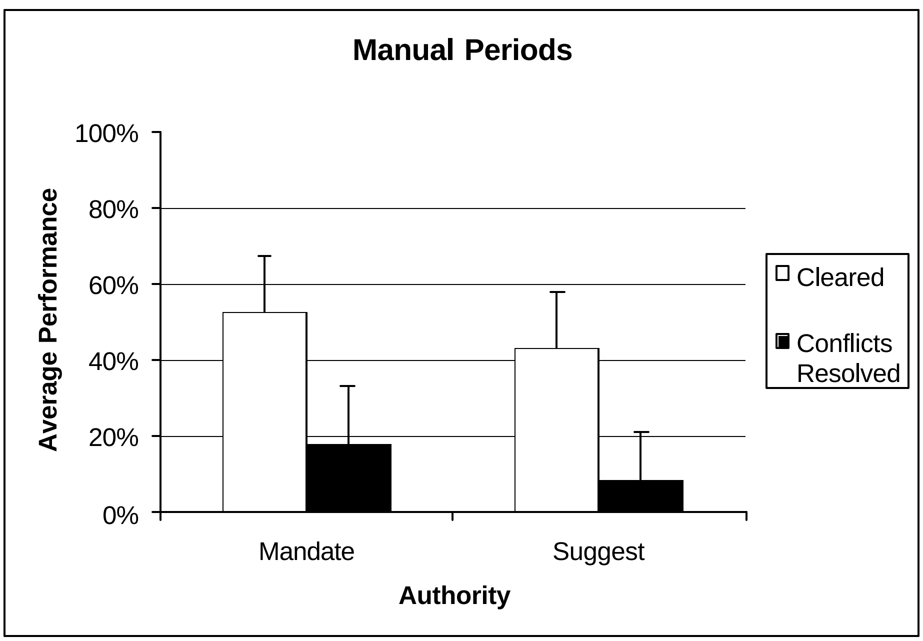


Figure 5. 4 Primary task performance during manual control by authority.

With respect to manual performance under the suggestion authority condition, a Shapiro-Wilks test on the conflicts resolved measure produced a significant result ($p<0.001$) and examination of a normal probability plot revealed a non-linear trend in the model residuals. As in the preceding analyses a square root transform was applied to the response, but it was not successful in correcting for the assumption violation. Consequently, a Kruskal-Wallis test was conducted on the manual performance data collected under the suggestion condition (the data set also included the manual control condition.). The effect of automation condition on manual performance was found to be significant under the suggestion setting of authority ($T(4)=13.59$, $p<0.009$). Additional applications of the Kruskal-Wallis test on pairs of automation settings indicated that subjects who experienced the manual control condition were able to resolve more conflicts than subjects experiencing manual control as part of AA applied to the information acquisition ($T(1)=7.27$, $p<0.007$), information analysis ($T(1)=6.93$, $p<0.009$) and action implementation ($T(1)=7.00$, $p<0.008$) functions. Another separate test revealed that subjects who experienced automated decision making also resolved more conflicts during periods of manual control than those who experienced automated information acquisition ($T(1)=4.54$, $p<0.033$). In general, these results suggest that there may be a carry-over effect of type of automation as part of AA on periods of manual control. An ANOVA on the manual performance data collected under the mandate authority condition did not reveal any significant results.

It is possible that the general nature of the primary task influenced the significant performance effects revealed through the resolved conflict data. The procedures for clearing targets in the Multitask simulation made it fairly difficult to resolve conflicts during the first minute of each trial. Although a conflict could occur anywhere on the display by one aircraft overtaking another, conflicts were more common near the center of the display where the eight possible flight trajectories intersected. It required as much as two minutes of travel time for slower aircraft to reach the center of the display. In that portion of the display, aircraft had more time to overtake one another and were brought closer together by the narrowing distance between the various trajectories. In fact, the average percent conflict resolved for all subjects during the first minute of task performance was 5.84% while it was 18.36% overall. All subjects, therefore, were more likely to have a relatively poor performance observation for the first minute of each trial, which (based on the procedures of the experiment) was always under manual control. Every additional minute of manual control after that provided an opportunity for subjects to improve their overall conflicts resolved score. The more manual minutes each subject experienced during a trial, the more opportunity there was to achieve a better conflict resolution score. An analysis of the number of minutes each subject spent under automated assistance revealed significantly more minutes under automation when it was suggested than when it was mandated ($F(1,48)=18.33$, $p<0.0001$). The difference observed in manual performance under the mandate versus suggest authority conditions for this measure is likely due in part to the subjects operating under suggestions spending less time using manual control and, therefore, receiving lower scores.

5.1.2 Secondary Task Performance

Objective workload data measured based on performance in the secondary task was organized into the same four types of data sets as used to make comparisons of performance in the primary task. This included data collected during periods of automated Multitask performance, manual Multitask performance, and manual control of the primary task under the mandate authority condition or the suggestion authority condition.

An ANOVA on performance in the gauge-monitoring task during automation of the primary task did not reveal significant effects due to automation type or invocation authority. The same result occurred for secondary task data collected during manual performance of the primary task, in general, and during manual control under mandated and suggested authority, in specific. However, for both automated ($F(4,48)=3.55$, $p=0.013$) and manual ($F(4,48)=2.72$, $p=0.040$) periods there were significant effects attributable to individual differences. Figure 5.5 summarizes the mean secondary task performance based on hit-to-signal ratio by automation type and automated and manual periods.

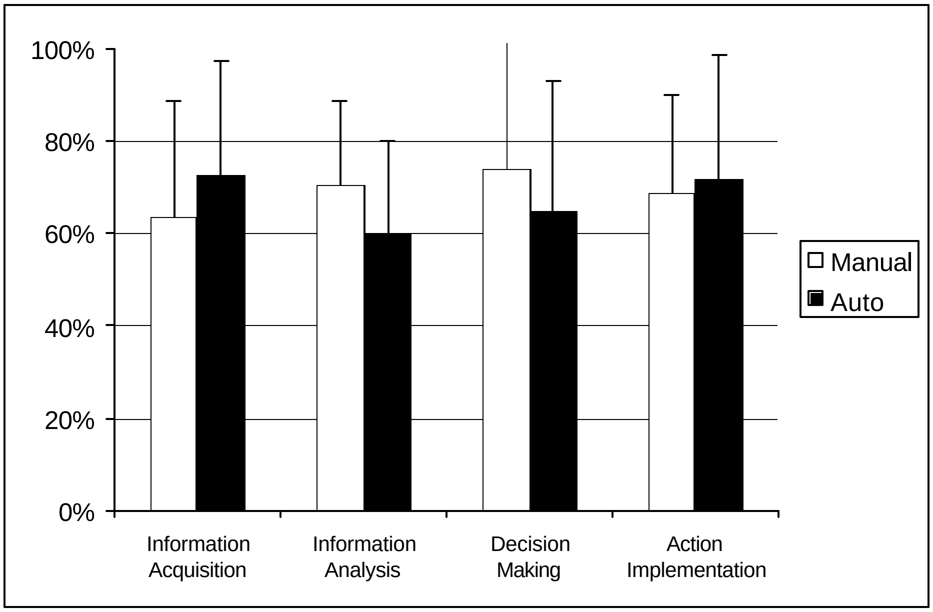


Figure 5. 5 Secondary task performance during manual and automated control.

5.2 Subjective Workload Ratings

The subjective workload measures were captured at the end of each trial, unlike the minute-by-minute recording of the objective workload measure. Consequently, two additional data sets were developed to analyze the subjective ratings of mental and temporal demand. The first set included 64 observations from both test trials for each subject assigned to the four automation conditions (4 automation types × 8 subjects × 2 trials). An ANOVA on this data did not indicate any significant effects of automation type of the form of authority on subjective workload.

The second data set included all ratings of workload collected during the experiment, including ratings by subjects assigned to the manual control condition (80 observations; 5 automation types × 8 subjects × 2 trials). An ANOVA on this data indicated a trend in temporal workload due to the type of automation ($F(9,70)=2.01$, p=0.051) with manual control receiving the highest workload rating. However, this finding was not statistically significant based on the alpha criterion for the study. Individual differences for both workload measures also produced significant effects (temporal: F(5,70)=2.88, p=0.020; mental: F(5,70)=2.73, p=0.026). In general, the results on these subjective measures of workload corroborated the results on secondary task performance.

5.3 Correlation Analysis

Correlation analyses using Pearson Product-Moment coefficients indicated significant linear associations between mental and temporal demand factors (r=0.5786, p<0.0001), mental workload and hit-to-signal ratio (r=-0.3194, p=0.004), and cleared targets and resolved conflicts (r=0.3364, p=0.002).

A comparison of primary and secondary task performance did not indicate that performance in the primary task could be explained in terms of operator workload, as influenced by the form of computer assistance. In addition, there did not appear to be a linear association between primary task performance and the subjective assessments of workload. However, the Pearson Product Moment coefficients did indicate that the percentage of cleared targets was predictive of conflict resolutions, as one would expect. The linear association between mental workload and the objective workload measure also indicated that subject perceptions of workload were consistent with workload observations based on secondary task performance. Finally, there was a highly significant correlation between the subjective assessments of mental workload and temporal demand, which indicates that the presentation of the dual-task scenario influenced these two measures in a similar fashion and that the VAS was as effective for recording changes in perceived time pressure as the MCH was for capturing changes in overall cognitive load.

6 Discussion

6.1 Interaction

The results of this study provide evidence that the effectiveness of AA is dependent upon both the type of automation presented to an operator and the type of invocation authority designed into the system. This is consistent with the findings of previous studies on the effects of automation type on individual performance (Endsley and Kaber, 1999; Kaber et al., 2002) and the performance effects of different forms of invocation authority over dynamic function allocations in an adaptive system (Hilburn et al., 1993; Kaber and Riley, 1999). However, contrary to the hypothesis that there would be a negative performance effect when computer suggestions were combined with AA of higher cognitive functions (information analysis and decision making), there was no evidence of an interaction effect of the two experimental manipulations. It is possible that the additional displays presented as part of the suggestion authority condition did not pose a consistent distraction to subjects and significantly degrade primary task performance. Furthermore, the highly significant effects due to individual differences in the secondary task, particularly when subjects were provided with automated assistance, may have overshadowed any effects the displays may have had on the workload measure.

6.2 Automation

During periods of automation assistance, subjects exposed to AA of information acquisition outperformed those experiencing AA applied to the decision making function. This differs from the results of Kaber et al. (2002) who observed the highest performance under AA of action implementation and minimal differences in primary task performance between AA of information acquisition and decision making. However, the results of these two studies do not entirely disagree. Based on graphical analysis of the current data on automation performance, the trend was in line with the stated expectations for the results, specifically that performance would be worse when AA was applied to higher-level functions. This observation is consistent with Kaber and Riley (1999), who found that operators might be hindered when computer assistance is provided in performance of higher-level cognitive functions. The expectations of this study were actually supported by the statistics in that there was a significant difference between information acquisition and decision making. The results obtained here are also in line with the findings of Kaber et al. (2002) in that there was no significant difference between the information analysis and decision making functions. However, this difference between information and decision automation did not appear to carry-over into the periods of manual control as part of AA. During these periods, performance in the primary task was similar across the various automation conditions.

It is possible that the superior performance under action implementation found by Kaber et al., as compared with the current study results may have been due to a difference in the physical layout of the task interface controls. Kaber et al. incorporated a separate numeric keypad for controlling Multitask and larger keys on the far side of a keyboard for the gauge. In the current experiment, the controls were closer together on a single keyboard and less discrete in nature, which may have led to keypunch errors when manually directing the portal. In general, subjects used the same hand to control the gauge and portal movement, which required frequent repositioning of the left hand on the right side of the keyboard. These keypunch errors were reported frequently by subjects in debriefings following the experiment.

6.3 Invocation Authority

In general, primary task performance improved during periods of automated assistance when the human was in charge of invoking automation (computer suggestion). This result is consistent with the hypothesis that human-initiated invocations of DFAs would improve primary task performance over system-initiated invocations under all automation conditions. However, this was not the case during periods of manual control. Although there were no significant effects due to invocation authority, the data suggests that performance in the primary task improves during manual control periods when the computer is in charge of controlling allocations.

In general, subjects spent significantly more time under automation when it was suggested compared to when it was mandated. Once subjects received automated assistance they rarely chose to return to fully manual control. It is likely that this was due to subjects generally believing that manual control of Multitask was more difficult than control with some form of automation assistance. In debriefings at the conclusion of the experiment, a substantial majority of subjects stated that they found the tasks easier when assisted by automation. This is consistent with the results of Kaber and Riley (1999) who found that subjects who received suggestions rather than mandates were less likely to return to manual control. Of the 32 subjects presented with some form of AA in the present study, only four agreed to a DFA under suggested authority while receiving automated assistance.

It is possible that by choosing to remain in an automated control mode, subjects may have voluntarily constrained the potential benefits of AA, including balancing operator workload and facilitating SA and effective performance when under manual control (Kaber et al., 2002; Hilburn et al., 1993). By ignoring the computer suggestions to return to manual control, subjects may have negated any potential carry-over effects of automation on manual performance and vice versa. In general, the results of this study show that although human operators may be effective at determining when to engage automation, they do not often optimize overall system performance by returning to manual control. Of course, this implies that there may be situations in which a computer mandate for a return to manual control may be more effective in terms of optimizing performance. The results also suggest that a blended form of authority, involving humans engaging automation and computer mandates of returns to manual control, may improve the potential for performance (,situation awareness) and workload benefits of AA in a complex system.

6.4 Primary Task

Taken together, the results on automation and invocation authority further underscore the potential pitfalls of implementing automated systems. The fact that automation is available does not necessarily mean that it is beneficial. The automation literature refers to difficulties in maintaining SA (Kaber and Endsley, in press), and operator complacency and skill decay (Parasuraman, 2000) as the potential problems of extended periods of automation, especially in the event of an automation failure. A human operator's judgment may also be more effective than automation, especially in novel situations that were not considered during the design of the automation (Inagaki, 2000). Specific applications of automation may present their own unique complications. For example, the Multitask simulation often presents scenarios in which a human operator can outperform the computer. In the case of information acquisition, the human operator may be able to identify more direct search paths for locating targets and the decision making automation does not always recommend an optimal strategy. Despite that in the majority of trials subjects believed their performance benefited from automation, this was not always the case.

6.5 Secondary Task

The results of performance in the secondary task did not indicate any significant differences in workload among the automation types or between the types of authority. Although gauge monitoring performance was fairly consistent throughout the experiment, the minute-by-minute observations on this measure varied enough to cause DFAs for all subjects. In a similar study using Multitask and the gauge monitor as an objective measure of workload, Kaber et al. (2002) found that the hit-to-signal ratio in the secondary task indicated higher workload when AA was applied to the information analysis function of the primary task, as compared to when it was applied to information acquisition or action implementation. A comparison of the means from the current study (Figure 5.5) suggests similar results. During periods of automation when AA was applied to information analysis, there appeared to be a decrease in secondary task performance (increase in workload) as compared to the other automation types. AA of decision making appeared to have similar results. It is possible that the lack of significance in this measure was due to the use of varying strategies by different subjects. In support of this, there were consistent significant individual differences in secondary task performance. The analysis of the subjective workload measures produced similar results.

A graphical comparison of mean secondary task performance during periods of automated and manual control for each automation type also suggests that the two AA types that induced the highest potential visual and/or cognitive load (information analysis and decision making) led to higher workload during periods of automated assistance. Further, the data suggest that AA of information acquisition and action implementation reduced workload as compared with periods of manual control. This implies that some forms of automation can potentially increase workload. This observation is in line with the findings of previous research (Selcon, 1991) demonstrating increases in pilot perceived workload due to the use of an automated decision aid in comparison to flight without decision automation.

The secondary task that served as an objective measure of workload in this experiment is analogous to peripheral responsibilities of an air traffic controller, such as radio communication, referencing general information (i.e., weather), and managing flight strip information. Air traffic controllers must attend to these tasks in conjunction with managing aircraft separation using the radar display. Like the gauge monitoring task, aspects of performance in any of these activities, such as reaction time, throughout the course of the task could potentially be used to provide an objective measure of workload.

Depending on the specific cognitive requirements of peripheral tasks, the actual activities of an air traffic controller may be more sensitive to changes in the role of automation in primary traffic management functions. Future research should investigate the use of peripheral ATC activities as embedded secondary task measures of workload as a basis for facilitating DFAs.

7 Conclusions

7.1 Caveats

The goal of this study was to assess the performance and workload effects of applying AA to different types of human-machine system information processing functions (based on a theoretical model of automation presented in the literature) and facilitating DFAs through two types of computer authority. Although the study had several significant results, their generalizability for the design of future ATC systems may be limited by a few factors. First, the population included naïve subjects. Participants received only two hours of training in advance of the experiment, so it is likely that seasoned air traffic controllers would exhibit different behaviors when presented with similar tasks as a result of their experience. Second, although the goals and functions of Multitask and the gauge monitor were analogous to actual complex systems, they represent a low-fidelity simulation of an ATC system. Real-world ATC includes numerous dimensions that do not exist in Multitask, such as aircraft altitude, multiple non-linear flight paths, and additional cognitive tasks such as managing flight strips and responding to pilot inquiries. Including these in the primary task and integrating peripheral responsibilities for the secondary task that are not as sensitive as gauge monitoring to individual differences might increase the applicability of additional research using these programs. Furthermore, the real-world stress associated with clearing for landing a commercial airliner containing hundreds of people is difficult to simulate in a laboratory experiment. Future research with the goal of improving ATC systems should aim to develop higher-fidelity systems and, when possible, include the participation of expert operators in order to provide information that would be directly applicable for the design of future systems.

7.2 Adaptive Automation in Air Traffic Control

It is expected that future additions to ATC systems will introduce new forms of automation (Duley and Parasuraman, 1999; Billings, 1997). If designed and implemented effectively, these new systems have the potential to both increase performance and decrease workload of future air traffic controllers. The challenge facing the designer is to ensure that automation does not impose any unanticipated burden on the operator in the form of information analysis and decision aids that are even more difficult to use than the original system, or computer assistance that places the operator in the role of monitor instead of an active decision maker. When designing an adaptive system, it is imperative that the designer understand the impact the automation will have on the human operator. Through empirical research, the type of automation, the locus of control, and the type of trigger of DFAs have all been shown to influence the effectiveness of AA in complex systems. These are all important features of automation that must be considered in system design.

7.3 Future Research

Adaptive automation represents a growing field that offers numerous directions for future research. The present study could therefore be expanded to investigate other factors significant to this area. For example, the literature review stressed the importance of maintaining SA in complex systems in order to encourage operator acceptance of workload and to optimize performance during periods of automated and manual control.

However, AA research directly examining SA remains fairly limited (e.g., Kaber and Endsley, in press). Measures of SA could be incorporated with the existing Multitask simulation in order to quantify the SA effects of AA of different types of human-machine information processing functions, various DFA authority conditions, or combinations of both. Situation awareness could also be investigated as a potential AA trigger, offering another option to the use of an objective, primary task workload measure, such as a secondary task.

Beyond the use of SA as a DFA trigger, it may be possible to make improvements to the existing model of the secondary task used in this work. A more embedded secondary task that does not require the operator to divide attention between two displays may add to the fidelity of the study. It is also possible that a secondary task that is less visually demanding may integrate differently with the primary task. Options include reaction time measures for responsibilities related to the primary task, such as time to respond to a request from a pilot.

Individual differences produced significant results in terms of operator workload, measured both objectively and subjectively, in this study ($p<0.05$). This may be due to the variety of ways in which subjects divided their attention across the two tasks. Steps could be taken to improve the learning that occurs during training and provide an opportunity to optimize attention distribution across the dual-task scenario and to reduce individual differences. This would potentially allow for a more robust assessment of automation type and the authority condition, as part of AA. In this study, participants were asked to describe their strategies in the debriefing following the last experimental trial. The subjects who performed best in both tasks developed strategies that allowed them to divide their attention across the tasks using time-phased features in the interfaces. These subjects learned that there were display spaces in-between the fixed flight trajectories for all aircraft where a vehicle would never appear. This allowed them to direct attention to the secondary task at regular intervals, as the portal moved through these regions. These strategies could be presented during experimental training sessions in order to reduce the variation in performance in a subject pool.

In this study, the subjects were not made aware of the true nature of the secondary task as an objective measure of workload and a trigger for AA of the primary task. It is possible that an advance understanding of this trigger could bias the operator into manipulating secondary task performance in order to influence the appearance of automation in the primary task. This is particularly significant if the operator perceives any advantage to automated or manual control. It is also likely that if a secondary task were integrated in a real-world system that the operators would eventually become aware of its significance in the overall system. Understanding how this knowledge would influence operator behavior in an AA system would therefore be an important addition to this area of research.

8 References

Barnes, M., & Grossman, J. (1985). The intelligent assistant concept for electronic warfare systems (Technical Report NWC TP 5585). China Lake, CA: Naval Warfare Center.

Billings, C. E. (1997). *Aviation automation: The search for a human-centered approach*. Mahwah, NJ: Erlbaum.

Bolstad, C., & Endsley, M. R. (2000). Shared displays and team performance. In D. B. Kaber and M. R. Endsley (eds.), *Human Performance, Situation Awareness and Automation: User-centered Design for the New Millenium* (pp. 311-317). Madison, WI: Omnipress.

Bubb-Lewis, C., & Scerbo, M. W. (1997). Getting to know you: Human-computer communication in adaptive automation. In M. Mouloua, & M. Koonce (Eds.), *Human-automation interaction: Research and practice* (pp. 92-99). Mahwah, NJ: Erlbaum.

Conover, W. J., (1980). *Practical Nonparametric Statistics (2nd ed.)*. John Wiley & Sons.

Duley, J. A., & Parasuraman, R. (1999). Adaptive information management in future air traffic control. In M. Scerbo & M. Mouloua (Eds.), *Automation technology and human performance: Current research and trends* (pp. 86-90). Mahwah, NJ: Erlbaum.

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics, 42(3)*, 462-492.

Golikov, Y. Y., & Kostin, A. N. (1997). The human-factors approaches to automation and allocation of functions between man and machine. In M. Mouloua, & M. Koonce (Eds.), *Human-automation interaction: Research and practice* (pp. 76-83). Mahwah, NJ: Erlbaum.

Hancock, P. A., & Chignell, M. H. (1993). Adaptive function allocation by intelligent interfaces. *Proceedings of the international workshop on intelligent interfaces* (pp. 227-229). Orlando, FL: ACM.

Harris, W. C., Hancock, P. A., & Arthur, E. J. (1993). The effect of taskload projection on automation use, performance, and workload. In R. S. Jensen and D. Neumeister (Eds.), *Proceedings of the seventh international symposium on aviation psychology* (pp. 178-184). Columbus, OH: The Department of Aviation, The Aviation Psychology Laboratory, The Ohio State University.

Hilburn, B., Molloy, R. Wong, D., & Parasuraman, R. (1993). Operator versus computer control of adaptive automation. In R. S. Jensen and D. Neumeister (Eds.), *Proceedings of the seventh international symposium on aviation psychology* (pp. 161-166). Columbus, OH: The Department of Aviation, The Aviation Psychology Laboratory, The Ohio State University.

Inagaki, T. (2000). Situation-adaptive autonomy: Trading control of authority in human-machine systems. In M. W. Scerbo and M. Mouloua (Eds.), *Automation technology and human performance: Current research and trends* (pp. 154-158). Mahwah, NJ: Erlbaum.

Kaber, D. B., & Endsley, M. R. (2001). *The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task*. Submitted to Theoretical Issues in Ergonomics Science.

Kaber, D. B., & Riley, J. M. (1999). Adaptive automation of a dynamic control task based on workload assessment through a secondary monitoring task. In M. W. Scerbo and M. Mouloua (Eds.), *Automation technology and human performance: Current research and trends* (pp. 129-133). Mahwah, NJ: Erlbaum.

Kaber, D. B., Wright, M. C., & Clamann, M. P. (2002). Human response to adaptive automation of early sensory/information acquisition functions and later stages of information processing in complex system operations (Final Report: NASA Langley Research Center Grant #NAG-1-01039). Hampton, VA: NASA Langley Research Center.

Moray, N., Inagaki, T., & Itoh, M. (2000). Adaptive automation, trust, and self-confidence in fault management of time-critical tasks. *Journal of Experimental Psychology, 6(1),* 44-58.

Parasuraman, R. (2000). Designing automation for human use: Empirical studies and quantitative models. *Ergonomics, 43(7),* 931-951.

Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model for types and levels of human interaction with automation. In *IEEE Transactions on systems, man, and cybernetics – Part A: Systems and humans*, 30, 286-297.

Prinzel, L. J., Scerbo, M. W., Freeman, F. G., & Mikulka, P. J. (1995). A bio-cybernetic system for adaptive automation. *Proceedings of the 39th annual meeting of the human factors and ergonomics society* (pp. 1365-1369). Santa Monica, CA.

Sanders, M., & McCormick, E. (1993) *Human factors in engineering and design* (7th ed.). McGraw-Hill.

Scallen, S. F., & Hancock, P. A. (2001). Implementing adaptive function allocation. *International journal of aviation psychology, 11(2)*, 197-221.

Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance, theory and applications* (pp. 37-63). Mahwah, NJ: Erlbaum.

Scerbo, M. W., Freeman F. G., Mikulka P. J., Parasuraman, R., Di Nocero F., & Prinzel, L. J. (2001). The efficacy of psychophysiological measures for implementing adaptive technology (Technical Report NASA/TP-2001-211018). Hampton, VA: National Aeronautics and Space Administration, Langley Research Center.

Scott, W. B. (February 1, 1999). Automatic GCAS: 'You can't fly any lower'. *Aviation Week & Space Technology, 150(5),* 76-79.

Selcon, S. J. (1990). Decision support in the cockpit. Probably a good thing? *Proceedings of the 34th annual meeting of the Human Factors and Ergonomics Society* (pp. 46-50). Santa Monica, CA: Human Factors and Ergonomics Society.

Sheridan, T. (1992). *Telerobotics, automation, and supervisory control*. Cambridge: MIT Press.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance (3rd ed.)*. Upper Saddle River, NJ: Prentice-Hall.

Wierwille, W.W., & Casali, J.G. (1983). A validated rating scale for global mental workload measurement application. *Proceedings of the Human Factors Society 27th Annual Meeting* (pp. 129-133). Santa Monica, CA: Human Factors Society.

Appendices

Appendix A: Multitask Hierarchical Task Analysis

The following high-level hierarchical task analysis (HTA) summarizes the tasks required to process aircraft (issue clearances) in Multitask. To the right of the HTA is the condition that automates each subtask, if applicable. A subtask that is automated does not need to be completed by the operator.

Task/Subtask	Automation
0. clear aircraft	
1. search	-
1.1 move portal	information acquisition
1.2 track aircraft	information acquisition
1.3 analyze aircraft	information analysis
1.4 prioritize aircraft	decision making
2. establish link	-
3. issue clearance	-
3.1 wait for link	action implementation
3.2 select aircraft	action implementation

Plan 0: 1
when 1 prioritizes aircraft and as required – 2 - 1 or 3 -1

Plan 1: 1.1
when aircraft found – 1.3 – 1.4
as required - 1.2

Plan 3: 3.1
when link is established – 3.2

Description
This model assumes expert performance. Errors such as incorrect selection of aircraft (e.g., issuing a clearance when no communication link has been established) are not considered.

1. Clear aircraft - The high level goal in Multitask is to maximize the number of aircraft cleared. In order to do this, the operator must search for aircraft, establish communication links with aircraft, and issue clearances to aircraft. Each of these three steps depends on the successful completion of the prior step (i.e., a clearance can not be issued unless a link is established and an aircraft is found).
2. Search – The search task requires the operator to locate and gather information about aircraft on the display.
   2.1 Move portal - Operators search for aircraft using the portal controlled with numeric keypad in horizontal, vertical, and diagonal directions.

2.2 Track aircraft - Once an aircraft is found, the operator must continue to follow the aircraft with the portal until the remaining steps are completed. If the operator fails to do so, he or she must return to Step 1.1.

2.3 Analyze aircraft – The operator gathers and attempts to memorize information about aircraft found, including:

- Type (military, commercial, or private)
- Stage of Processing (new aircraft, communication link established, clearance issued, or in transition)
- Location (direction of travel, distance to due date, proximity to other aircraft)

Note: when recalling proximity to other aircraft it is also necessary to recall the locations of other aircraft.

2.4 Prioritize aircraft – In order to prioritize an aircraft, the operator must recall the information from Task 1.3 (analyze aircraft). The methodology used by the operator to determine whether or not to aircraft is dependent on the individual's strategy. The only aircraft property that can restrict the operator's progress is the stage of processing. Only new (red) aircraft can be selected for establishing a communication link. Only linked (yellow) aircraft can be issued a clearance. The other relevant properties of aircraft than can be used for prioritization are:

- Speed (based on type)
- Relative location (distance to due date and other aircraft)

If an aircraft is determined not to be a priority, the operator must restart the search. The result of this task is the identification of a single aircraft.

3. Establish link – once an aircraft is identified, the operator starts the process of establishing a communication link by left-clicking the aircraft with the mouse controller. Once that is completed, the operator returns to Task 1.
   Note: Upon completion of this step, although the operator must return to task 1, the subtasks can be resolved very quickly if the operator continues to process the same aircraft he or she established the link for.
4. Issue clearance – The issue clearance task represents the final processing stage for each aircraft.

4.1 Wait for link - Since the stage of processing restricts the issuance of a clearance, the operator must wait seven seconds after completing Task 2 (establish link) before a clearance can be issued.

4.2 Select aircraft – The operator's last subtask is to begin to issue a clearance by right-clicking the aircraft with the mouse controller. Once this subtask is completed, the operator returns to Task 1.

Appendix B: Instructions for Participants

I. Introduction

Welcome and thank you again for volunteering to participate in this experiment. This investigation is being conducted to study the effects of authority over automation in a complex system of different aspects of information processing on your performance in a computerized "radar" monitoring and aircraft-clearing task as well as a gauge-monitoring task. The former task will require you to use a standard PC keyboard and a 2-D mouse linked to a cursor in processing graphic targets presented on a display screen. The gauge-monitoring task will require you to maintain a moveable pointer within an "acceptable" range on a fixed scale presented on a display by using the keyboard. Different levels and durations of computer assistance will be provided in performing the tasks.

The procedures to be followed include:

1. an equipment familiarization period,
2. a 15-minute training session during which you will be provided with instructions on manual radar monitoring performance,
3. a 15-minute training session and instructions on automated task performance,
4. a five-minute training session and instructions on the gauge-monitoring task,
5. an explanation of the contest,
6. a 24-minute practice session requiring you to perform the aircraft clearing and gauge-monitoring tasks while switching between the manual and automated modes of the former task,
7. a 10-minute rest period, and

8. a two minute demonstration of how to switch modes of automation in the aircraft processing task
9. testing in two 20-minute trials separated by a 10-minute break.

Your experimental participation will take place on a single day over approximately two and a half hours. You are asked to complete all experimental testing; however, you may discontinue your participation at any time.

All experiment instructions will be presented to you verbally. If you do not understand certain instructions, you will be able to ask questions after completion of each step in the procedure. You will also have an opportunity to ask questions about the experiment during training and the rest periods, which will occur between each training session and experimental trial. You must follow all instructions given before and during the experiment carefully. If you fail to follow instructions or the equipment malfunctions, I will stop testing. In these cases, you will be allowed to read the task instructions and ask any questions you may have, or I will correct the system malfunction. Once your full comprehension of the instructions is ensured and the equipment is in working order, testing will re-commence.

Informed Consent and Data Sheets

[Give the subject the informed consent form. Summarize the informed consent for the subject and encourage them to read the form.]

This form summarizes the information that has been presented to you thus far and identifies the persons responsible for the study. The form also addresses University liability to the experiment. I encourage you to read the form. This form will not be associated with any of the other survey forms used in this experiment. In order to participate in this study you must have 20/20 or corrected vision. Potential risks include: (1) soreness of the hand and forearm muscles from extensive use of a mouse in controlling the Multitask© simulation and use of a keyboard in controlling the gauge-monitoring task; and (2) visual strain and/or fatigue in viewing the simulation displays through a conventional personal computer (PC) monitor. These risks are not substantially different from those associated with my everyday PC use. In the event that you experience fatigue or discomfort, please inform me immediately.

Please sign and date this form.

[Make a copy of the signed form for the subject to keep]

[Present the subject with the data sheet.]

This form asks about your personal characteristics and will serve to verify your qualifications for the study. Please take a few moments to complete the survey. If you have any questions, I will be happy to address them. This form, like the informed consent form, will not be associated with any of the other survey forms used in this experiment to record your performance on the various tasks.

[Have subjects complete all payment forms for participation. Be sure to record the time and date that the subject started participation.]

This is the payment form that will be used to calculate your compensation for participating in this experiment. Please fill out the information.

II. Familiarization

The experimental equipment is comprised of two displays, a full-size keyboard, and a mouse controller, which have been set up here to resemble a typical working scenario in a "radar" monitoring operation. The displays are high-resolution graphics monitors commonly used with PC-based computing systems. During training and actual testing the room will be darkened to eliminate any glare reflected off the surfaces of the displays. The keyboard is a standard 124 key programmable unit. You will only be required to use the control and shift keys, the spacebar, the numeric keys '1' through '9', and the 'Enter' key on the numeric portion of the keyboard **[Point to the keys that will be used]**. The mouse is a simple controller facilitating movement of the cursor on the "radar" monitor in both the horizontal and vertical directions.

This completes a brief overview of the testing equipment you will use. Do you have any questions concerning the setup?

III. Multitask Training

[Have subject sit in chair in front of task displays. Run Multitask, the Gauge Monitor, and PowerPoint slides with the target images.]

[Start a MultiTask practice trial with the following parameters as a demonstration:]

One target

No automation

No portal

The “radar” monitoring task involves clearing randomly appearing graphic targets in the form of small representations of aircraft. Aircraft are represented by three different graphic icons depicting military, commercial, and private aircraft.

[Show the subject the three different icons]

Military targets travel the fastest, followed by commercial, then private. The targets will follow one of the eight approach paths represented on the scope by yellow lines. The goal is to clear targets for landing before they reach the airport at the center of the display. Clearing, or processing, targets involves two steps: (1) establishing a communication link and (2) issuing a clearance for landing. When targets first appear on the display, they are red. You will establish a communication link on red targets by pointing the cursor with the mouse controller and clicking on the desired target with the **left** mouse button. Once the target has been clicked, it will flash for a few seconds, and then turn yellow. To issue a clearance, point the cursor with the mouse controller and click on the desired target with the **right** mouse button. The target will flash for a longer period of time, and then turn green. Once the target turns green, the clearance has been issued and no other interaction is necessary. Green targets can no longer be affected by mouse clicks, nor can they interfere with other targets on the display. The green target will continue to move toward the due date for landing.

[Demonstrate the procedure to the subject. Allow the aircraft to hit the due date. When finished, restore the portal.]

During the practice and test trials, most of the ‘radar’ display will be blacked out. A small portion of the display will be made visible by a ‘portal,’ which acts as a movable spotlight. The direction of the portal’s movement can be controlled with the numeric portion of the keyboard. ‘4’ and ‘6’ move the portal right and left, respectively, while ‘8’ and ‘2’ control up and down motion. The portal can also be moved diagonally by pressing ‘1’, ‘3’, ‘7’, or ‘9’. You should use the portal to find targets on the display. Once a target is exposed with the portal, you may clear it by using the left and right mouse buttons as described earlier.

Your selection of the various targets will be dependent upon which target you feel requires the most immediate attention. All targets should be cleared before they reach the airport at the center of the screen. However, there will be several targets on the display at one time, all of which will move at different speeds and directions. There is a strong possibility that targets will collide with each other if not cleared quickly. Target rewards are given when a target turns green (target is fully cleared). An additional reward is given if the cleared target was on a collision course with another target. It is unlikely that you will be able to clear all the targets on the display.

[Ask the subject if they have any questions]

We will now begin the first of four training sessions.

I ask at this time that you remove your watch or timepiece from your person and place it in a pocket. Due to the nature of this experiment, keeping your attention focused on the aircraft-clearing task is very important.

[Select ‘Manual Practice’ trial, the number of targets, and ‘None’ for the automation level for training in the Multitask application. Enter subject number, random seed, and a trial number. Do not check the box for ‘Run Gauge.’]

You will have 15 minutes to practice clearing targets on the “radar” scope. You should take advantage of this practice period to formulate a conscious strategy for effectively clearing targets. You may take time to learn the controls and get a feel for the interface, as your performance in this training period will not be recorded.

Feel free to move the keyboards and mouse around to a comfortable position during training.

If you encounter a difficulty during the session please do not hesitate to bring it to the attention of the experimenter, who will immediately assist you.

Please commence testing by depressing the 'Enter' key on the keypad on my mark.

[Once training is complete, offer an optional 2-min rest period.]

If you need to use the rest room or get a drink, feel free to take a 2-minute break at this time.

Unless you have any questions, we will begin the second of the four training sessions.

[Select 'Automation Practice' trial, the number of targets, and the appropriate automation level for training in the Multitask© application. Enter subject number, random seed, and a trial number.]

You will now be provided additional training in a manner similar to the previous exercise; however, in this session, the computer will provide some assistance in locating targets. As in the first practice session, your task will be to clear the randomly appearing targets on the "radar" display following the various approach paths before they collide or expire.

[Read instructions from the Level of Automation Definition document appropriate for the automation group to which subject has been assigned.]

You will have 15 minutes to practice clearing targets on the "radar" scope. You should take advantage of this practice period to formulate a conscious strategy for effectively clearing targets.

Please commence testing by depressing the 'Enter' key of the keypad on my mark.

[Once training is complete, allow for a 2-min rest period.]

There will be a 2-minute break prior to beginning the next practice period. If you need to use the rest room or get a drink, please do so at this time.

IV. Gauge-Monitor Training

You will now be provided training in the gauge-monitoring task.

[Instruct subject to shift gaze to gauge monitoring display.]

The gauge-monitoring task involves maintaining a moving, white pointer within an "acceptable" (green) range on a fixed scale display. The pointer moves "up" and "down" the scale in a random manner.

You will maintain the pointer in the "acceptable" range by depressing the 'Control' or 'Shift' keys on the right side of the keyboard.

Your selection of these keys will be dependent upon whether the pointer moves into the "upper-unacceptable" (red) range or "lower-unacceptable" range on the scale display. The 'Shift' key should be depressed when the pointer moves into the "lower-unacceptable" range causing the arrow to move upward. Similarly, the 'Control' key should be depressed when the pointer moves into the "upper-unacceptable" range causing the arrow to move downward.

Note that the ‘Control’ and ‘Shift’ control keys **do not** function when the pointer is within the “acceptable” range.

Depending upon the timing of your control actions, ‘Correct Detection’ of an unacceptable pointer deviation will be recorded.

Unless you have any questions, we will begin the third of the four training sessions.

[Select ‘Training (5 min)’ trial in the gauge-monitoring package. Enter subject number, and random seed. Move the mouse to a corner of the screen.]

You will now have five-minutes to practice maintaining the moving pointer within the ‘acceptable’ (green) range on the fixed scale display.

Please commence testing by depressing the ‘Enter’ key of the keyboard on my mark.

[Once training is complete, allow for a 2-min rest period.] There will be a 2 min break prior to beginning the next practice period. If you need to use the rest room or get a drink, please do so at this time.

V. Contest Rules

The participant who attains the highest combined score in both tasks across both experiment trials is eligible for a $30 gift certificate to a local restaurant. This form summarizes the rules for the contest.

[Present the form, Contest Rules and Score Sheet, to the participant]

The information on this form includes the rules of the contest and instructions for redeeming the prize. It also serves as a verification of your score. In order to redeem your prize, you must present this form to the experimenter. If you wish to participate, please sign and date the form at this time. I will enter your score in the blanks when you complete the experimental trials. If you have any questions about the contest, I can answer them now.

VI. Dual-Task Training

You will now be provided additional training in a manner similar to the previous exercises; however, this time we will combine the target clearing and gauge-monitoring tasks. We will also combine the automated computer assistance with the manual target-clearing task. (Note that control of the gauge-monitoring task will remain unchanged.)

As in the other practice sessions, your tasks will be to maintain the moving pointer in the "acceptable" range on the fixed scale display and to clear the randomly appearing aircraft on the "radar" display following the various approach paths before they collide or expire from the display.

Unless you have any questions, we will begin the training session.

[Select 'Dual-Task Practice' trial, number of targets, and the appropriate automation level for training in the Multitask application. Enter subject number, random seed, and a trial number. Make sure the check box to 'Run Gauge' is checked. Do not enter performance and standard deviation data.]

You will have 24 minutes to practice eliminating targets on the "radar" scope and maintaining the moving pointer within the "acceptable" (green) range on the fixed scale display. You should focus your attention on both tasks and take advantage of this practice period to formulate a conscious strategy for effectively clearing targets while maintaining the pointer in the acceptable range.

At periodic intervals, the computer will switch between ‘Manual’ and ‘Automated’ modes of the target-clearing task. During those intervals, the computer assistance for locating targets will either be turned on or off.

Please commence testing by depressing the ‘Enter’ key on my mark.

[During training, Multitask will toggle between manual and automated modes once every two minutes. Allow for a 10-min rest period.]

This completes the training sessions. There will be a 10-minute break prior to beginning the testing. If you need to use the rest room or get drink, please do so at this time.

[During the break, import the data from GPRAC for the subject into Excel or Access and calculate the mean and standard deviation for the last 16 minutes. Record these figures.]

VII. Authority Demonstration

[Run Multitask for 2 min Authority Training for the subject’s scheduled condition]

You will now be given a demonstration on how to approve automation during the experimental trial. In order to receive automated assistance for clearing the targets, you and the computer need to agree that it is needed. The status of the automation is presented in this box **[Point to the authority status window]**. This box shows you (1) the current state of the automation, on or off, and (2) the current state of recommendations for automation. The computer has an icon for its recommendations and you have an icon for your approvals.

[Read the following for computer suggest]

For this trial, the computer must recommend automation or manual control before you can approve it. If the computer believes automation is necessary, you will receive a message from the computer that automation is 'recommended,' you will hear a beep, and the computer icon will turn **yellow**. To request automation, you can press the space bar, your icon will turn **yellow** and automated assistance will start. If the computer believes manual control is necessary, you will see the 'recommended' message, you will hear a beep, and the computer icon will turn **blue**. To accept manual control, press the space bar, your icon will turn **blue** and you will return to manual control.

VIII. Subjective Rating Forms

In order to assess the task workload that you experience during experimental testing, you will complete subjective comparisons of various demand factors.

[Show the two demand comparison form]

Following both test trials, you will be required to complete subjective ratings of perceived workload. You will rate mental workload using this form.

[Show hard copy of MCH]

You will complete this form by following the flow chart to the box that you feel best describes the level of mental workload you experienced in order to perform the target clearing task

[Show hard copy of temporal demand scale]

You will complete this form by drawing a straight vertical line on the scale directly below each of the factors indicating the level of temporal demand experienced during the target-clearing task. A definition of temporal demand as it applies to this experiment is at the top of the form. Please consider only the target-clearing task when rating the various demands.

IX. Testing

Experimental testing will occur in a manner similar to the last training exercise. As in the practice sessions, your tasks will be to maintain the moving pointer in the "acceptable" range on the fixed scale display and to clear the randomly appearing targets on the "radar" display following the various approach paths before they collide or expire.

Actual testing at the level of automation you experienced during the last training session will now occur for 20 minutes during which control will shift between manual and automated assistance at various times.

Upon completing the testing, you will be given another 10-minute break.

[Select 'Experiment' trial, number of targets, and level of automation for testing. Enter subject number and random seed. Enter the performance and standard deviation values captured during the last training session. Run the gauge.]

Unless you have any questions, we will begin testing.

Please commence testing by depressing the 'Enter' key on my mark.

[When the trial is completed, give the subject the subjective rating forms and allow for a 10 min rest period. Be sure to record the combined gauge and multitask scores for the contest.]

You will now be allowed an additional 10-min break to rest.

[Read the following list and return to beginning of Testing section to present next trial involving different automation schedule.]

We will now begin another trial.

If you encounter a difficulty during the session please do not hesitate to bring it to the attention of the experimenter, who will immediately assist you.

This completes your session. I'd like to thank you for your participation and I can now answer any questions you have about the experiment.

[Enter the time on the payment form and sign at the top. Enter combined score on contest sheet, if applicable.]

Level of Automation Definitions

Manual

You must move the portal (manually direct the radar) using the keys on the numeric keypad.

You must clear targets by pressing the left mouse button on red targets (establishing a communication link), then the right mouse button on the target once it turns yellow (issuing a clearance). It should be noted that in this mode, targets may be partially processed; however, no partial reward will be given. A target is either fully cleared or not cleared at all.

Information Acquisition

The portal is automatically directed. The portal moves around the display in a circular motion. The number keys do not affect its motion. You can also cause the portal to automatically track a target once an aircraft is revealed on the display. When tracking is on, the portal will remain over the target, keeping it visible until the target is fully processed by the user (i.e., the target turns green). To activate or turn off automated tracking, press the ‘Enter’ key. No additional aids are provided.

You must clear targets by pressing the left mouse button on red targets (establishing a communication link), then the right mouse button on the target once it turns yellow (issuing a clearance). It should be noted that in this mode, targets might be partially processed; however, no partial reward will be given. A target is either fully cleared or not cleared at all.

Information Analysis

You must move the portal (manually direct the radar) using the keys on the numeric keypad.

A decision aid **[Show Decision Aid Image]** appears on the left side of the display summarizing the current layout of the targets, their speeds and sectors of airspace in which an aircraft may be involved in a potential collision. The **Icon** column shows the heading and classification for the target (commercial, military, or private aircraft), as well as the stage of processing. The **Conflict** column alerts the subject if the target may be involved in a collision. If a conflict is detected, the first character shows the position (in the form of a compass direction) of the target, while the second character shows the position of the other aircraft involved. The **Speed** column presents the target's speed in knots, while the **Distance** shows the distance to the due date.

You must clear targets by pressing the left mouse button on red targets (establishing a communication link), then the right mouse button on the target once it turns yellow (issuing a clearance). It should be noted that in this mode, targets may be partially processed; however, no partial reward will be given. A target is either fully cleared or not cleared at all.

Decision Making

You must move the portal (manually direct the radar) using the keys on the numeric keypad.

A decision aid **[Show Decision Aid Image]** appears on the left side of the display summarizing the current layout of the targets and sectors of airspace in which the targets are traveling. The **Icon** column shows the heading and classification for the target (commercial, military, or private aircraft), as well as the stage of processing. The **Position** column gives the heading of the targets as a compass direction. The list of targets in the decision aid is sorted so the aircraft with the highest processing importance are at the top of the decision aid. Target importance is based on whether it is involved in a conflict and its distance to the due date. In this mode, you can use the decision aid as a basis for processing targets.

You must clear targets by pressing the left mouse button on red targets (establishing a communication link), then the right mouse button on the target once it turns yellow (issuing a clearance). It should be noted that in this mode, targets may be partially processed; however, no partial reward will be given. A target is either fully cleared or not cleared at all.

Action Implementation

You must move the portal (manually direct the radar) using the keys on the numeric keypad.

A decision aid **[Show Decision Aid Image]** appears on the left side of the display summarizing the current layout of the targets and sectors of airspace in which the targets are traveling. The **Icon** column contains a symbol, which shows the stage of processing for the target.

You must clear targets by pressing the left mouse button on red targets (establishing a communication link). After the communication link is established, the clearance will be issued automatically, so there is no need for a second mouse click. It should be noted that in this mode, targets might be partially processed; however, no partial reward will be given. A target is either fully cleared or not cleared at all.

SUBJECTIVE RATING OF PERCEIVED TEMPORAL DEMAND

Indicate the level of temporal demand experienced during the target elimination task for each of these factors by drawing a straight vertical line on the scale directly below. How much time pressure did you feel due to the rate at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Temporal Demand

Low High

SUBJECTIVE RATING OF PERCEIVED MENTAL WORKLOAD