The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task

Article *in* Theoretical Issues in Ergonomics Science · March 2004

DOI: 10.1080/1463922021000054335

2 authors:

David B Kaber
University of Florida

Mica R. Endsley
SA Technologies

The Effects of Level of Automation and Adaptive Automation on Human Performance, Situation Awareness and Workload in a Dynamic Control Task

DAVID B. KABER[1] AND MICA R. ENDSLEY[2]
[1]Department of Industrial Engineering, North Carolina State University, Raleigh, NC 27695-7906, USA
[2]SA Technologies Inc., 4731 East Forest Peak, Marietta, GA 30066, USA

Keywords: Level of automation (LOA); adaptive automation (AA); human-centered automation; situation awareness; mental workload; embedded-secondary task

This paper extends previous research on two approaches to human-centered automation: (1) intermediate levels of automation (LOAs) for maintaining operator involvement in complex systems control and facilitating situation awareness; and (2) adaptive automation (AA) for managing operator workload through dynamic control allocations between the human and machine over time. Some empirical research has been conducted to examine LOA and AA independently with the objective of detailing a theory of human-centered automation. Unfortunately, no previous work has studied the interaction of these two approaches, nor has any research attempted to systematically determine which LOAs should be used in adaptive systems and how certain types of dynamic function allocations should be scheduled over time. In the present research, we briefly review the theory of human-centered automation and LOA and AA approaches. Building on this background, we present an initial study that attempts to address the conjuncture of these two approaches to human-centered automation. An experiment was conducted in which a dual-task scenario was used to assess the performance, SA and workload effects of low, intermediate and high LOAs, which were dynamically allocated (as part of an AA strategy) during manual system control for various cycle times comprising 20%, 40% and 60% of task time. The LOA and automation allocation cycle time (AACT) combinations were compared to completely manual control and fully automated control of a dynamic control task performed in conjunction with an embedded secondary monitoring task. Results revealed LOA to be the driving factor in determining primary task performance and SA. Low-level automation produced superior performance and intermediate LOAs facilitated higher SA, but this was not associated with improved performance or reduced workload. The AACT was the driving factor in perceptions of primary task workload and secondary task performance. When a greater percentage of primary task time was automated, operator perceptual resources were freed-up and monitoring performance on the secondary task improved. Longer automation cycle times than have previously been studied may have benefits for overall human-machine system performance. The combined effect of LOA and AA on all measures did not appear to be "additive" in nature. That is, the LOA producing the best performance (low level automation) did not do so at the AACT, which produced superior performance (maximum cycle time). In general, the results are supportive of intermediate LOAs and AA as approaches to human-centered automation, but each appears to provide different benefits to human-machine system performance. This work expands the current understanding of these approaches and provides additional information for a developing theory of human-centered automation.

1. Introduction

At this point in time, serious problems associated with automation and numerous human-machine system errors have been documented (cf., Billings, 1991; Moray, 1986; Sarter & Woods, 1995; Wiener & Curry, 1980). These problems have been associated with various deficiencies in human operator states, including vigilance decrements, complacency and loss of situation awareness (SA), which have also been discussed at length in numerous studies (cf., Carmody & Gluckman, 1993; Endsley, 1987; Endsley & Kiris, 1995; Parasuraman, Mouloua, Molloy & Hilburn, 1993; Parasuraman and Riley, 1997; Wiener, 1988). In general, a key underlying factor that has emerged as a contributor to human performance problems in complex, automated systems control is human out-of-the-loop (OOTL) performance (see Kessel and Wickens (1982), and Young (1969)).

Out of the loop performance problems are characterized by a decreased ability of the human operator to intervene in system control loops and assume manual control when needed in overseeing automated systems. First, human operators acting as monitors have problems in detecting system errors and performing tasks manually in the event of automation failures (Billings, 1988; Wickens, 1992; Wiener & Curry, 1980). In addition, they have a more complex system to monitor. In a review of automation problems, Billings (1988) noted six major aircraft accidents that could be traced directly to failures in monitoring automated systems or the flight parameters controlled by the automated systems.

In addition to delays in detecting that a problem has occurred necessitating intervention, operators may require a significant period of time to reorient themselves to the current state of the system after a failure and develop a sufficient understanding of the state in order to act appropriately. This delay may prohibit operators from carrying out the very tasks they are required to perform or diminish the effectiveness of actions taken. Wickens and Kessel (1979, 1981) conducted laboratory studies demonstrating longer system recovery times and poor response accuracies for operators who had been removed from control loops in advance of critical events requiring intervention.

These two types of SA problems (failure to detect and failure to understand the problem) have been hypothesized to occur through three major mechanisms:

(1) changes in vigilance and complacency associated with monitoring;
(2) assumption of a passive role instead of an active role in controlling the system; and
(3) changes in the quality or form of feedback provided to the human operator (Endsley & Kiris, 1995).

Each of these factors can contribute to the OOTL performance problem. In addition, automated systems, by nature of their complexity, also challenge higher levels of SA (comprehension and projection) during ongoing system operations. The general idea here is that certain LOAs may lead to OOTL performance and loss of SA. This issue was explored by assessing the impact of a broad range of LOAs on operator SA in our experiment; however, the adaptive nature of contemporary automated systems was also considered.

To overcome some of the ills created by historically technology-centered approaches to automation (systems which automate whatever can be automated) a philosophy of human-centered automation has been proposed (Billings, 1991, 1997). Billings (1997, p. 4) defined human-centered automation as facilitating a cooperative relationship in the control and management of a complex system with potential benefits for performance. Sheridan (1997) said that human-centered automation has many alternative meanings ranging from "allocate to the human the tasks best suited to the human, allocate to the automation the tasks best suited to it," through "achieve the best combination of human and automatic control, where 'best' is defined by explicit system objectives." The meanings he presented span from a function-oriented perspective to a mission-oriented view.

The goal of human-centered automation is to create systems that retain the human operator in control loops with meaningful and well-designed tasks that operators are capable of performing well in order to optimize overall human-machine system functioning. Billings (1997, p. 4) said that human-centered automation should ensure that automation does not leave the human with a fragmented and difficult job. It should define the assignment of tasks to a human and computer in controlling an automated system such that a team effort is achieved (Billings, 1997; Endsley, 1996). High levels of human-machine system performance may be achieved through human-centered automation by ensuring that the human has the capability to monitor the system, that they receive adequate feedback on the state of the system, and that the automation functions in predictable ways (Billings, 1997, p. 39), all of which support achievement of SA. Our intention here is not to attempt to fully define how to create human-centered automated systems (see Billings (1997) for a review), but rather to explore two separate research thrusts which have been proposed as at least partial methods for achieving automation designs which meet the goal of human-centered automation.

Several approaches have been proposed that challenge the traditional division of human-automation task responsibility in complex systems, specifically automation of as many tasks as possible and assignment of the human to the role of monitor. These approaches redefine the assignment of functions to people and automation in terms of a more integrated team approach. Two orthogonal and possibly complementary approaches can be defined along the axes of Figure 1. One approach seeks to optimize the assignment of control between the human and automated system by keeping both involved in system operations. This has been labeled Level of Automation (LOA) or "level of control" (see Draper (1995)). The other recognizes that control must pass back and forth between the human and the automation over time depending upon situational demands, and seeks to find ways of exploiting this understanding to increase human performance. This has been labeled Adaptive Automation (AA) or Dynamic Function Allocation (DFA) (cf., Corso & Moloney, 1996). One objective of this section is to show how these approaches may be effective for achieving human-centered automation.

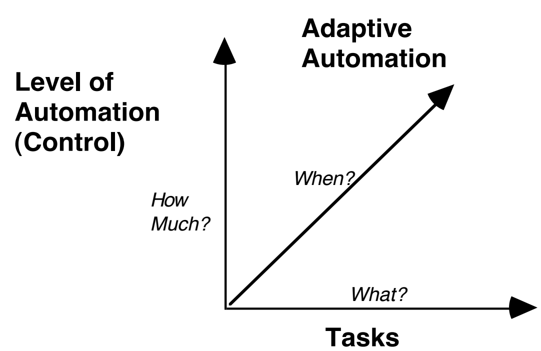

Figure 1. Approaches to human-centered automation (from Endsley, 1996).

Level of automation refers to the level of task planning and performance interaction maintained between a human operator and computer in controlling a complex system (Billings, 1991; Kaber, 1997). Sheridan (1997) discussed various degrees of automation that were defined in terms of the autonomy of complex system information sensing and control execution. These degrees dictated the extent to which a human was involved in system control and the level of computer task aiding. The LOA approach defines the assignment of system control between a human and computer in terms of the degree to which both are involved in system operations (Endsley, 1996). It emphasizes the interaction between a human operator and computer. The objective of this approach is to find LOAs that are best suited to human capabilities and capacities (the general objective of human-centered automation). Billings (1991) also defined various LOAs and was concerned with how automation might perform some portion of activities for (or interact with) humans and aid them in a course of action. He defined LOAs in terms of the autonomy of functions an operator would typically control and the capabilities of a system to perform response execution and to monitor its own action.

Rouse (1988) conceptualized AA as varying degrees of computer assistance in complex systems control based on the nature of a situation, including task characteristics and the state of the human operator. He proposed structuring human-automation interaction on the basis of "what" is to be automated, "when" a task is automated, and "how" it is automated. Through empirical study, Rouse (1977) identified several advantages of AA including support of human performance, dynamic definition of a coherent task role for operators, and the capability to maintain acceptable human workload levels in system control. Parasuraman et al. (1992) said that AA represents an optimal integration of automation with the human operator based on the level of operator workload. They placed an emphasis on the state of the operator. Scerbo (1996) summarized that under AA, different types of automation may be initiated and terminated dynamically based on situational demands placed on the system, inclusive of the operator. Kaber and Riley (1999) provided a contemporary definition of AA saying that it concerns the scheduling of the allocation of control between a human operator and computer over time, with the intent of improving human performance as part of complex systems operations or maintaining operator involvement in system control in order to reduce OOTL performance problems.

The key difference between the AA and LOA approaches is that AA involves dynamic control allocations (automated or manual, varying over time), and LOA involves static function assignments (Kaber, 1997; Parasuraman, Sheridan, & Wickens, 2000), defining the degree to which a task is automated. Later we provide a perspective on AA that has yet to be empirically studied, specifically dynamic allocations of a broad range of LOAs during task performance. We will first briefly review relevant research that has been conducted on the LOA & AA approaches and then present our experiment, which seeks to integrate them by exploring the decision space shown in Figure 1.

1.1. *Level of Automation as an Approach to Human-Centered Automation*

Automation does not exist in an all or none fashion. Rather it can be applied to different aspects of a task in varying degrees, creating different levels of task autonomy (Endsley & Kaber, 1999). A number of different taxonomies, or hierarchies of LOA have been developed. In an early seminal work, Sheridan and Verplank (1978) developed a hierarchy of LOAs in the context of undersea teleoperator control, as shown in

Table 1. Sheridan and Verplank's (1978) hierarchy of LOAs.

(1) Human does the whole job up to the point of turning it over to the computer to implement,
(2) Computer helps by determining the options,
(3) Computer helps to determine options and suggests one, which human need not follow,
(4) Computer selects action and human may or may not do it,
(5) Computer selects action and implements it if human approves,
(6) Computer selects action, informs human in plenty of time to stop it,
(7) Computer does whole job and necessarily tells human what it did,
(8) Computer does whole job and tells human what it did only if human explicitly asks,
(9) Computer does whole job and decides what the human should be told,
(10) Computer does the whole job if it decides it should be done, and if so, tells human, if it decides that the human should be told.

Table 1. This hierarchy includes varying allocations for determining options and selecting among them. The LOAs were differentiated in terms of decision making and action selection functions. In addition, the list focuses on what the human should be told by the system (i.e. issues of information display as well as task functioning). Sheridan & Verplank's objective was to define "who" (the human or computer) has control in a more definitive sense and not to explicitly describe how an operator and automation might share core information processing functions in complex system control.

Endsley (1987) developed a LOA hierarchy in the context of the use of expert systems to supplement human decision making. This hierarchy stipulated that a task could be performed using:

(1) manual control — with no assistance from the system;
(2) decision support — by the operator with input in the form of recommendations provided by the system;
(3) consentual artificial intelligence (AI) — by the system with the consent of the operator required to carry out actions;
(4) monitored AI — by the system to be automatically implemented unless vetoed by the operator; and
(5) full automation with no operator interaction.

This list is most applicable to cognitive tasks in which operator ability to respond to, and make decisions based on, system information (with expert system assistance) is critical to overall performance. Ntuen and Park (1988) developed a similar five level taxonomy of automation in the context of a teleoperation system. Both of these taxonomies can be seen to be similar to selected levels of the Sheridan and Verplank hierarchy, however, a completely manual level is also considered. The lowest LOA in Sheridan and Verplank's hierarchy represented what has subsequently been labeled direct teleoperation (Draper, 1995) and may also involve telerobot control if the human turns the task over to automation.

Building on this work, Endsley and Kaber (Endsley & Kaber, 1997; Endsley & Kaber, 1999) developed a ten-level taxonomy of LOA to provide wider applicability to a range of cognitive and psychomotor tasks requiring real-time control within numerous domains including air traffic control, aircraft piloting, advanced manufacturing and teleoperations. All of these domains have many features in common, including (1) multiple competing goals, (2) multiple tasks competing for an operator's attention, each with different relevance to system goals, and (3) high task demands under limited time resources.

Four generic functions intrinsic to these domains were identified that form the basis for this taxonomy:

(1) monitoring - which include taking in all information relevant to perceive system status (e.g., scanning visual displays);
(2) generating - formulating options or task strategies for achieving goals;
(3) selecting - deciding on a particular option or strategy; and
(4) implementing - carrying out the chosen option through control actions at an interface.

Ten LOAs were then systematically formulated by assigning these functions to the human or computer or a combination of the two, as shown in Table 2.

Below are high-level descriptions of the various LOAs in the taxonomy intended to introduce them. Some of these are expanded later in the discussion of the specific experimental task and conditions, including specification of human and machine activities (performance) at each level.

Table 2. Endsley and Kaber's (1999) LOA taxonomy for human-computer performance in dynamic, multitask scenarios.

	ROLES			
LEVEL OF AUTOMATION	MONITORING	GENERATING	SELECTING	IMPLEMENTING
1. Manual Control	Human	Human	Human	Human
2. Action Support	Human/Computer	Human	Human	Human/Computer
3. Batch Processing	Human/Computer	Human	Human	Computer
4. Shared Control	Human/Computer	Human/Computer	Human	Human/Computer
5. Decision Support	Human/Computer	Human/Computer	Human	Computer
6. Blended Decision Making	Human/Computer	Human/Computer	Human/Computer	Computer
7. Rigid System	Human/Computer	Computer	Human	Computer
8. Automated Decision Making	Human/Computer	Human/Computer	Computer	Computer
9. Supervisory Control	Human/Computer	Computer	Computer	Computer
10. Full Automation	Computer	Computer	Computer	Computer

(1) Manual —The human performs all tasks including monitoring the state of the system, generating performance options, selecting the option to perform (decision making) and physically implementing it.

(2) Action Support — At this level, the system assists the operator with performance of the selected action, although some human control actions are required. A teleoperation system involving manipulator slaving based on human master input is a common example.

(3) Batch Processing —Although the human generates and selects the options to be performed, they then are turned over to the system to be carried out automatically. The automation is, therefore, primarily in terms of physical implementation of tasks. Many systems, which operate at this fairly low level of automation, exist, such as batch

processing systems in manufacturing operations or cruise control on a car.

(4) Shared Control — Both the human and the computer generate possible decision options. The human still retains full control over the selection of which option to implement, however, carrying out the actions is shared between the human and the system.

(5) Decision Support —The computer generates a list of decision options, which the human can select from, or the operator may generate his or her own options. Once the human has selected an option, it is turned over to the computer to implement. This level is representative of many expert systems or decision support systems that provide option guidance, which the human operator may use or ignore in performing a task. This level is indicative of a decision support system that is capable of also carrying out tasks, while the previous level (shared control) is indicative of one that is not.

(6) Blended Decision Making — At this level, the computer generates a list of decision options, which it selects from and carries out if the human consents. The human may approve of the computer's selected option or select one from among those generated by the computer or the operator. The computer will then carry out the selected action. This level represents a high-level decision support system that is capable of selecting among alternatives as well as implementing the selected option.

(7) Rigid System —This level is representative of a system that presents only a limited set of actions to the operator. The operator's role is to select from among this set. He or she cannot generate any other options. This system is, therefore, fairly rigid in allowing the operator little discretion over options. It will fully implement the selected actions, however.

(8) Automated Decision Making — At this level, the system selects the best option to implement and carries out that action, based upon a list of alternatives it generates (augmented by alternatives suggested by the human operator). This system, therefore, automates decision making in addition to the generation of options (as with decision support systems).

(9) Supervisory Control — At this level the system generates options, selects the option to implement and carries out that action. The human mainly monitors the system and intervenes if necessary. Intervention places the human in the role of making a different option selection (from those generated by the computer or one generated by the operator); thus, effectively shifting to the Decision Support LOA. This level is representative of a typical supervisory control system in which human monitoring and intervention, when needed, is expected in conjunction with a highly automated system.

(10) Full Automation — At this level, the system carries out all actions. The human is completely out of the control loop and cannot intervene. This level is representative of a fully automated system where human processing is not deemed necessary.

It should be noted that the taxonomy represents a wide range of feasible assignments of the four functions of system(s) monitoring, and options generation, selection and implementation to human, computer and human/computer combinations. There may well be other combinations of these four core functions that were not specifically listed in the taxonomy; however, these were not deemed to be either technically or practically feasible (although certainly not ruled out). We also stipulated that the order of LOAs presented was not necessarily ordinal on any factor, but rather was a preliminary assessment of a possible order. This is a key difference between this taxonomy and the historical hierarchies of LOAs.

Endsley & Kaber's LOA taxonomy provided several advantages over the previous ones in that it identified numerous LOA combinations not included in the other taxonomies. By representing how system monitoring, process planning, decision making and response execution can all be assigned to a human operator or computer, or shared between the two, the taxonomy also provides greater detail on "who" (the human or computer) is doing "what" at each LOA, as compared to the historical hierarchies of degrees of automation. Furthermore, the present list does not focus only on decision making and defining authority. The functions it is based upon are also generic enough to be applicable to a wide variety of domains and task types. Most importantly, by systematically allocating the basic task components to the human or computer or a combination of the two, Endsley and Kaber's taxonomy provided the key advantage of allowing a careful empirical assessment of which aspects of automation might be helpful or harmful to human performance in conjunction with that system.

More recently, Parasuraman, Sheridan and Wickens (2000) provided a similar LOA taxonomy that considered whether each of four functions was automated, including (1) information acquisition, (2) information analysis, (3) decision selection, and (4) action implementation. The information acquisition, decision selection and action implementation factors of this model are identical to the monitoring, selection and implementation features of the Endsley and Kaber taxonomy. This model does not however explicitly consider the option generation (planning) function. Instead it provides information analysis as a separate function to be automated. Included in this aspect of automation are the integration of information into a single value, advanced displays such as prediction displays or emergent perceptual feature displays, and information displays that present context-dependant summaries of the data to the user. While certainly important for human performance, it remains open to debate whether displays constitute a type of automation. The Parasuraman et al. (2000) model does not explicitly provide LOAs, but rather stipulates that each of the four factors can be automated at differing levels.

While these taxonomies have many similarities and some significant differences, more importantly for the design of systems is empirical research that provides designers with some indication of the expected effects of automation design options presented by considering LOAs on subsequent human-system performance. Unfortunately, until recently, this type of work has been lacking.

How to specify the "best" LOA does not turn out to be as straight-forward as one might think, however. For particular applications, Parasuraman et al. (2000) list human performance, automation reliability, and cost associated with outcomes as appropriate criteria for selection of an LOA for a particular application. However, in order to effectively use these LOA decision criteria, there must be data available on system performance or an elaborate iterative design process must be undertaken involving specification and evaluation of both types and levels of automation in order to make decisions about which LOAs may be optimal for supporting system performance. In order to develop broader design guidance and theory on how to develop human-centered automation, however, we are interested in the effect of LOA choices on human cognitive processing and performance, independent of particular applications.

When the system is functioning well, a determination of whether automation is better than manual control is really a function of how good the automation algorithm is for that particular application, and the expertise of the individual human operators (level of experience, skills, etc.) It is easy to see that even a fairly modest system may outperform unskilled or inexperienced operators, but a much more complex and capable system may be needed to outperform a highly skilled operator. A similar issue comes into play in comparing performance at intermediate LOAs, which are comprised of both human performance and automation performance. From a research standpoint, such comparisons are very limited in use because they are almost completely specific to the particular application and situation and, thus, such findings do not generalize well to other systems.

A more useful comparison of LOAs can be found in examining how well the human-machine system performs, not just when the automation can address a specific situation, but also when it cannot. The ability of the human operator to detect and take over under system failure (either through a breakdown, or through reaching particular conditions for which it is not programmed) forms the crux of the OOTL error problem and thus should be central to decisions regarding LOA choices. Situation awareness, workload and trust or confidence level are often measured as relevant indices of human performance under such circumstances as well as normal conditions. In addition, the performance of the human in the face of confusing system input (e.g. mode awareness errors (see Sarter and Woods (1995)), or incorrect system input, should be explicitly considered in determining the effects of LOA choices on human performance.

Some empirical research has been conducted that specifically examines these issues across multiple LOAs. Endsley and Kiris (1995) conducted a study involving a simulated automobile navigation task and found that SA was lowest under full automation and only partially lower under the intermediate LOAs as compared to fully manual task performance. They also demonstrated that this lower SA correlated with OOTL performance decrements when the automated aid failed and operators were forced to perform the task manually. Of the three sources of OOTL decrement posited — vigilance decrements, poor feedback under automated control, or passive vs. active processing — they empirically identified passive processing as the only likely culprit in this study. Intermediate LOAs were found to have value in reducing the OOTL deficit, as compared to full automation.

Endsley and Kaber (1997, 1999) proceeded to more fully explore this problem space in the context of their 10-level LOA taxonomy. They used a dynamic control task, Multitask© (Kaber & Endsley, 1995), which is an abstract simulation of a radar-monitoring task. It involves multiple simulated targets, which must be eliminated prior to reaching their expiration time or colliding with one another. Each target has different reward points associated with its elimination and penalty points for failures. Multiple targets compete for operator attention simultaneously; therefore, the operators must develop a complex strategy to maximize performance.

During test trials, subjects experienced simulated automation failures, or shifts from the LOA, to which they had been assigned, to Manual Control. The failures occurred at random points in time and lasted a short duration. During other trials, simulation freezes occurred at random times in order to administer SA queries using the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1988). In this way, we examined the impact of simulated automation failures on OOTL performance problems and LOA effects on SA.

The results indicated that LOA significantly impacted both task performance and OOTL performance problems. By systematically comparing each LOA, several significant findings regarding the effects of automation could be determined. First, performance was enhanced by computer aiding in task implementation (LOAs 2-10; see Table 2); however, it was hindered at the LOAs involving joint human-computer option generation (LOAs 4-8). Automation providing aiding in the action selection aspect of the task (LOAs 6, 8 and 9) did not significantly affect system performance when compared to purely human decision making. We also found that performance at high LOAs (9 and 10) was better than Manual Control performance; however, it was never as good as when low-level automation was used to provide assistance in manual implementation aspects of the task, exclusively (LOA 2). (This can be attributed in part to limitations in the automation algorithm.)

Operator ability to recover from, and perform in the event of, automation failures was, however, superior under LOAs requiring some human interaction in task implementation. Automation strategies that allowed operators to focus on future task processing (through advanced queuing of tasks) led to increases in time-to-recover task control following an automation failure, and subsequent manual performance was worse. Therefore, this type of automation may contribute to OOTL performance decrements.

Situation awareness and workload were also impacted by LOA; however, lower operator workload and better Level 2 SA were found at some of the higher LOAs in the taxonomy (LOAs 6, 8, 9 and 10), contradicting findings of other research (cf., Endsley & Kiris, 1995) that better SA occurs at intermediate LOAs. However, the study did also find that Level 3 SA was better at intermediate LOAs. Endsley and Kiris (1995) observed no difference in perceived workload across LOAs. This discrepancy in results was viewed as a possible effect of the different experimental tasks used in each study and needs further exploration.

In other research, Kaber et al. (2000) conducted an experiment to examine the potential benefits of intermediary LOAs to human operator performance, SA and workload in using a high-fidelity simulation of a telerobot for nuclear materials handling. Kaber et al. (2000) studied six of the LOAs presented in Endsley and Kaber's (1999) taxonomy, including Manual Control (LOA 1), Action Support (LOA 2), Batch Processing (LOA 3), Decision Support (LOA 5), Supervisory Control (LOA 9) and Full Automation (LOA 10), representing the range of LOAs found to impact Multitask© performance. Results indicated that higher LOAs led to improved performance in terms of time-to-task completion and number of errors committed under normal conditions. Subjects also reported lower levels of workload at progressively higher LOAs. In general, the experiment revealed the benefit of computer implementation of the teleoperation task.

During simulated automation failures, however, subjects functioning at higher LOAs preceding the failure were slower in reacting to system malfunctions, they took longer in re-orienting themselves to correct the failure, and committed more errors when assuming Manual Control, as compared to subjects using intermediate LOAs prior to a failure. The intermediate LOAs permitted a greater degree of human involvement in the control loop during normal system operations, potentially increasing subject awareness of system status prior to a failure and promoting faster recovery times.

Situation Awareness Global Assessment Technique results revealed a significant impact of LOAs on Level 3 SA (projection of system states), with lower SA occurring under higher LOAs, as in Endsley and Kaber's (1999) study. Action Support (computer assistance in the implementation aspect of the task) again facilitated the highest level of SA, as compared to all other LOAs. This study confirmed many of the findings of Endsley and Kaber's (1999) research in a more realistic task, and the results have significant implications for the design of automation for telerobot control.

While not exploring the range of LOAs, Moray et al. (2000) recently reported on a study that implemented three LOAs included in Sheridan and Verplank's hierarchy (levels 5, 6 and 7 (see Table 1)), which roughly correspond to Endsley and Kaber's LOAs 5, 8 and 10. When system reliability was good (over 90%), human performance was fairly good (at LOA 5). When system reliability was poorer, however, human performance suffered. The performance of the machine in this case was still helpful in terms of carrying out actions, even though diagnoses were poor and more false plant shutdowns occurred. They concluded that different LOAs might be needed, depending on the time criticality of the tasks at hand.

Finally, Lorenz et al. (2001) compared three LOAs in an automated diagnosis system: a fault finding guide (low level support), a decision support tool (Endsley and Kaber's LOA 5), and a higher level aid (LOA 8). They found improved performance in participants working with both the medium and higher LOAs, as compared to the lower LOA group. Perceived workload was not different between the three LOA groups. Under automation failure, however, the medium LOA group performed the worst. This was attributed to differences in information sampling strategies that appeared to be induced by the differences in the three conditions. Under the medium LOA condition, participants were less likely to see information that was important for diagnosis. This study points to the fact that many issues, other than the LOA itself, can affect performance. Careful attention to the design of the system and the feedback provided to the operator are required under any LOA (see Norman (1989)).

In general, these studies are supportive of an approach to human-centered automation that features lower LOAs retaining operators in the control loop. The research demonstrates that even when full automation of a task may be technically possible, it may not be desirable if performance of the joint human-machine system is to be optimized. Intermediate LOAs may be preferable for certain types of tasks and certain system functions in order to maintain human operator SA at high levels by allowing them to perform critical functions and, at the same time, to moderate workload in comparison to that experienced under manual control.

1.2. *Adaptive Automation as an Approach to Human-Centered Automation*

As previously mentioned, AA has also been put forth as an approach to human-centered automation with the objective of reducing OOTL performance problems. The literature has defined a number of strategies to AA (see Scerbo (1996) for a thorough review), or for allocating system control between humans and computers, including:

(1) critical events – DFAs triggered by occurrence of events critically impacting system goals (e.g., malfunction) (Hilburn, Molloy, Wong & Parasuraman, 1993);
(2) performance measurement – DFAs triggered by degradations in human monitoring performance below a

criterion measure (Parasuraman, 1993);
(3) psychophysiological assessment – real-time assessment of operator workload (using for example physiological measures – electroencephalogram (EEG) signals or heart-rate variability) as basis for decision to automate (Pope, Comstock, Bartolome, Bogart, & Burdette, 1994; Byrne & Parasuraman, 1996); and
(4) behavior modeling – DFAs occur to human and computer to achieve predetermined pattern of overall system functioning (Rouse, Geddes & Curry, 1986).

Similar to the psychophysiological assessment strategy, Hancock and Chignell (1988) also proposed that a strategy for AA involving comparison of current and future states of operator workload as well as system performance would be a desirable a basis for DFAs.

Like LOA, assessments have been made of the effect of AA on human operator performance, SA and workload in complex systems control. In general, research has demonstrated that AA may be effective for certain complex system task types and when certain durations of DFAs are used. Early empirical research explored the behavior modeling and performance measurement strategies for managing operator workload and affecting performance by scheduling dynamic allocations of manual control and full automation during system performance (i.e., the work considered a binary approach to AA – only minimum and maximum levels of system automation were considered even though it is commonly accepted that automation may vary along a continuum and that there are many levels that could be allocated in an adaptive system). Specific issues examined by this work have included the optimal frequency and duration of cycles of automation and manual control, as well as who (the human or computer) has ultimate authority for managing DFAs over task time. Parasuraman (1993) studied a performance-based strategy using the Multi-Attribute flight Task (MAT) Battery. He required subjects to perform the tracking and fuel management tasks manually during 30-min trials while the systems monitoring tasks was statically or adaptively automated based on a model of system performance or how well a subject performed manual monitoring. Parasuraman (1993) found monitoring performance (failure detections) to improve with periodic allocations of manual control to operators (every 10 min for a duration of 10 min) during automated control, as compared to the static automation condition. The two AA strategies appeared to be equivalent in terms of performance benefits.

Hilburn et al. (1993) also used the MAT Battery to study an AA strategy involving either operator or computer managed control allocations. They found that performance in the tracking subtask of the MAT Battery was significantly better when the computer versus operator determined manual and automated control allocations. Further, AA, in general, was found to produce better monitoring performance than purely manual monitoring. Hilburn et al. (1993) stated that performance degradations observed during the experiment were due to operators frequently cycling between full automation and manual control and that excessively short-automation cycles of 2 min compounded this effect.

Contrary to these findings, Scallen et al. (1995) found significant improvements in tracking task performance when control shifted between manual and automated modes every 15 s as compared to every 60 s in piloting tasks. However, the dynamic function allocations in this study followed a predetermined schedule. The DFAs in

Hilburn et al. (1993) study were adaptive in nature and based on operator performance levels.

Negative performance consequences of AA may not be limited to excessively short control cycles, as long cycle AA has also been shown to cause inefficient operator performance. Hilburn et al. (1993) presented another experiment in which AA was applied to the monitoring aspect of the MAT Battery during 120-min trials with automated control being allocated every 10 min for a duration of 10 min. They found operator monitoring under AA to be extremely poor with only 32% of system malfunctions detected, as compared to completely manual monitoring, which produced a 75% detection rate. That is, AA did not help performance in comparison to manual monitoring. They stated that operators tended to place excessive trust in the automation and, when manual control was reallocated, their abilities were limited due in part to longer duration exposure to automation. The performance problems with long cycles may be attributable to the OOTL performance problem, in general, and to complacency and skill decay over extended periods, more specifically.

Similar to this study, Parasuraman et al. (1996) examined the effect of model- and performance-based AA on human monitoring performance using the MAT battery during long-duration trials (three sessions totaling 90 min). They used the tracking and fuel management tasks and developed an automated engine status task. Under the AA strategies, manual control of the engine status task was periodically allocated based on when monitoring performance was expected to be at its worst (model-based AA) or if individual monitoring performance in previous automated periods did not meet criterion levels (performance-based AA). Automation failure detection was compared across the AA conditions and a static automation condition. In general, their results revealed that both AA approaches enhanced monitoring performance over static automation during long duration tests. Taken together with the results of Hilburn et al. (1993) study, this research demonstrates that during extended trials AA may only be superior to manual control and static automation under certain task conditions.

Some limitations of these studies include a focus on the application of AA to psychomotor tasks, such as monitoring and tracking, and the binary perspective of AA (studying the allocation of manual control during fully automated operations and vice versa). More contemporary work has focused on AA applied to cognitive tasks, the use of the psychophysiological strategy to AA, and the issue of who has authority over dynamic control allocations. Hilburn et al. (1997) conducted a study in the context of Air Traffic Control (ATC) to examine whether a critical events strategy to adaptive allocation of strategic planning advisories could be used to reduce operator workload and optimize human performance. Experienced air traffic controllers were required to control an ATC simulation with or without the assistance of an automated tool for managing and controlling arrival traffic. The automation tool detected planning conflicts or projected separation conflicts and offered the human operator advice aimed at solving the detected conflicts. Hilburn et al. (1997) used three automation schemes including constant manual control, constant automation and the AA condition (under which automation was invoked only during high traffic conditions). They found that the AA condition resulted in the smallest increase in mental workload across trials.

Kaber and Riley (1999) explored a performance-based strategy to AA by using a secondary task measure of workload to facilitate control allocations in a complex dynamic control task (the primary task). They required subjects to perform the Multitask© simulation along with a simple, gauge-monitoring task during 10-min trials. Differences between secondary task performance (gauge monitoring) in the absence of the primary task, and gauge monitoring as part of the dual-task scenario, were observed and used as a basis for directing operator managed control allocations in the primary Multitask©. Adaptive automation involving shifts between manual control and partial automation (Blended Decision Making (LOA 6)) of the primary task were mandated for one group and merely suggested for another. Kaber & Riley found significantly improved manual, primary task performance and enhanced secondary task monitoring under automation for the mandated-AA group, with the opposite results occurring for non-mandated AA subjects. The average subject workload marginally exceeded an objectively established criterion by using the secondary task measure to direct AA of the primary task.

These studies demonstrate that the critical events and performance approaches to AA may be effective for moderating operator workload in various cognitive tasks. Unfortunately, contemporary studies have not examined the use of such AA strategies to prevent OOTL performance problems, including a loss of operator SA (Endsley & Kiris, 1995). Kaber and Riley's (1999) work also provided insight into who should decide whether and when automation should be invoked–the human or computer. Consistent with other research (see Scerbo (1996)), they demonstrated that humans might not be the best judges of DFAs. Additional empirical evidence is needed, however, in order to resolve this question in various contexts, including psychomotor and cognitive task performance. Moray et al. (2000) also examined the issue of ultimate decision making authority under AA of their process control simulation (an apartment complex heating system). Operators were posed with two tasks including monitoring and controlling a temperature gauge, and monitoring an automated fault-management system (diagnosis and control of leaks and breaks) that functioned at various levels of reliability (identified above). The temperature control task was always performed manually. With respect to decision making responsibility for DFAs, they offered that the characteristics of the situation are key and that there are circumstances (time- and safety-critical) in which overriding authority for control allocations must be given to automation. They demonstrated that under high-time stress conditions (a pipe "break"), operator response time and accuracy is worse and that high-level AA should be implemented for control and to diagnose whether human intervention is possible.

In general, both historical and contemporary AA research is supportive of a theory of human-centered automation and defining dynamic changes in control function allocations between humans and computers based on states of the collective human-machine system. This research demonstrates that AA may be superior to other forms of complex system control for certain task types and durations. Specifically, AA may provide performance benefits to operators involved in monitoring, psychomotor and dynamic control tasks. These benefits appear to result from maintaining operator involvement in active control and managing workload, which may serve to prevent OOTL performance problems including complacency, vigilance decrements, and a loss of SA and manual skills.

2. Direction of Current Study

In general, the results on the effectiveness of AA and LOA approaches to human-centered automation have been positive, specifically they have demonstrated both approaches to promote human-machine system performance, moderate operator workload, and facilitate SA. However, to this point in time, no work has considered how LOA and AA may interact to affect performance or SA. There is a need to examine the combined effectiveness of intermediate LOAs and adaptive allocation of LOAs during dynamic control tasks involving cognitive functions in order to define the role of each in human-centered automation in terms of performance, SA and workload. Hilburn et al. (1997) evaluated the impact of AA on cognitive function performance in the context of simulated ATC tasks, but automated assistance was either active or not and varying degrees of assistance were not considered. The high-level goal of the experiment as part of this research was to describe the relative effects of LOA and AA, as well as the interaction of these approaches, on human performance, SA and workload in a complex system control task.

More specific research needs motivating this work include the fact that all empirical studies of AA conducted thus far have been limited to the binary perspective of the concept, including full automation and manual control allocations (e.g., Scallen, Hancock, & Duley, 1995; Parasuraman, Mouloua, & Molloy, 1996; Hilburn, Jorna, Byrne, & Parasuraman, 1997). Few other LOAs have been examined in AA research, such as allocating supervisory control during manual functioning or allocating manual control during a batch processing mode. Furthermore, we only have limited knowledge regarding appropriate frequencies and durations of dynamic control allocations during experimental tasks. Previous research has evaluated that combination of frequency and duration of DFAs yielding a half-automated and half-manual task (Parasuraman, Mouloua, Molloy, & Hilburn, 1993; Hilburn, Molloy, Wong & Parasuraman, 1993).

For example, different durations of automation allocations (e.g., the Full Automation LOA (10)) have not been systematically examined while holding the overall frequency of manual control allocations fixed across task time. Excessively frequent cyclings between manual control and full automation (e.g., every 2 min) appear to cause deficits in system performance (Hilburn, Molloy, Wong, & Parasuraman, 1993). However, there also appear to be differences in the effectiveness of AA with extremely short cycles (less than 1 min) with the longer of these yielding performance improvements (Scallen, Hancock, & Duley, 1995). Long cycle times (e.g., 10 min), in general, have been shown to produce both decrements (Hilburn, Molloy, Wong, & Parasuraman, 1993) and improvements (Parasuraman, 1993; Parasuraman, Mouloua, & Molloy, 1996). It is possible that generally short-cycle times do not allow operators to adapt to interface changes in order to perform tasks effectively; whereas, long-cycle times may reduce the frequency of operator involvement in system operations (manual control) and possibly lead to OOTL performance problems, including complacency and vigilance decrements. It is likely that there are also workload and SA tradeoffs across different DFA schedules, although these have not been studied. Optimal cycle times remain to be defined.

The present research addressed these needs and current limitations in knowledge of AA. The work also involved determining which intermediate LOAs, included in Endsley and Kaber's (1999) taxonomy, and DFA schedules provide superior performance in a dynamic control task. One important question concerning the integration of these approaches is what LOAs should be adaptively allocated during complex tasks in order to enhance performance and SA? For example, will providing manual control opportunities to operators during supervisory control of a system improve overall performance beyond that previously observed with conventional AA? Furthermore, will different DFA durations, less or greater than 10 min, benefit performance while holding the frequency of allocations fixed?

At the outset of this research, we hypothesized that intermediate LOAs would support operator SA and that short AA control cycles would have a negative impact on performance. It is also possible that intermediate LOAs, distributing higher-level information processing functions between the human and computer, may improve performance as a result of the computer guidance positively influencing operator task planning. Beyond this, we speculated that lower levels of automation would support improved performance, based on Endsley and Kaber's (1999) previous results. Based on Endsley and Kiris' (1995) findings, we also expected that SA would degrade to a greater extent with AA cycles involving allocation of high-level automation as a result of operators being OOTL. With respect to the duration of AA cycles, we expected long cycles to yield improved performance and lower operator workload, but that SA might be degraded in comparison to shorter cycles.

These are specific postulates that were considered in the experiment by comparing model-based AA (control allocations of set durations and timing) with both manual control and full automation using the dual-task paradigm employed by Endsley & Kaber (1997, 1999) and Kaber & Riley (1999). (As in Parasuraman et al. (1996), a model-based approach to AA was used to achieve a predetermined cycle of manual and automated performance and to allow for examination of the performance effects of periodic task automation in a controlled manner.) Other hypotheses could be formulated regarding the interaction effects of LOA and AA. However, no previous research exists upon which to base such hypotheses, other than the previous research evaluating each approach independently.

This research is important because there is a need for empirical results that can be used as a basis to further detail a theory of human-centered automation for application across domains. This type of work is urgent in LOA and AA studies, as there currently exists no formal theoretical framework of research results for guiding the design of adaptive systems and intermediate LOAs together in order to effectively manage operator workload and facilitate SA.

3. Methodology

The specific objective of the experiment was to determine whether manual control allocations during system operations at various LOAs, or allocations of a broad range of LOAs during manual operations, benefit human-machine system performance, operator SA and workload in comparison to completely manual control and fully automated performance. Unlike in Endsley and Kaber's (1999) study, manual control allocations in this experiment were not described as automation failures, but rather opportunities for subjects to augment overall system performance, as might be the case in introducing human operators to implementations of complex real-world adaptive systems. It is unlikely that in making AA a reality that manual control periods as part of DFAs would be characterized as pseudo-automation failures, although from a conceptual standpoint this may be accurate.

3.1. *Subjects*

Thirty university students (13 males and 17 females) having 20/20, or corrected to normal visual acuity, and some personal computer (PC) experience participated in the experiment for monetary compensation. Subjects ranged in age from 18 to 40 years (mean = 21.67) and all, but two, were right handed.

3.2. *Tasks and Experimental Conditions*

The subjects performed modified versions of the dynamic control and secondary tasks used by Endsley and Kaber (1999) and Kaber and Riley (1999), including the Multitask© simulation and gauge-monitoring task. The Multitask© simulation was automated under a subset of the LOAs described in the Introduction and presented in Table 2, including Manual Control (LOA 1), Batch Processing (LOA 3), Shared Control (LOA 4), Blended Decision Making (LOA 6), Supervisory Control (LOA 9), and Full Automation (LOA 10). These LOAs represent a broad range of automation and were identified by Endsley and Kaber (1999) as being significantly different in terms of human performance of Multitask©.

3.2.1. *Multitask©*

The Multitask© simulation has been used successfully in several recent empirical studies (e.g., Bolstad & Endsley, 2000; Endsley & Kaber, 1997, 1999; Clamann, Wright & Kaber, 2002) to investigate, for example, the effectiveness of shared displays for facilitating SA in team operations, and has been validated as a robust paradigm for evaluating human-automation interaction (i.e., it is possible to discriminate among many theoretical LOAs in terms of performance, SA and workload using this task). Here we provide a detailed description of the characteristics of the task, the operator goal, and interface features and functionality.

The simulation presents targets (multiple tasks) to an operator in the form of square shapes of different sizes and colors on the mock radarscope. The targets travel at various speeds towards a processing deadline at the center of the display (see Figure 2). An operator's goal is to select and eliminate targets (i.e., carry out the tasks) by collapsing their areas before they reach the deadline or collide with one another. The specific methods by which target selection and elimination occur through the interface represent the LOAs in Table 2 and are described in detail below. As targets are collapsed, reward points (see upper-left corner of display) are added to a total score (see center of display). Penalty points are assessed for target expirations or collisions (also see upper-left corner of display) and deducted from the total score. The size (small, medium, large) and color (red, blue, green) of each target encodes reward and penalty points. The exact point values for each target are displayed as data tags attached to them. The high-level goals of an operator are to maximize reward points and minimize penalty points. The speed at which targets travel and their initial distances from the center deadline provide information on the time available for processing a task, and are considered to be factors in target selection. The total travel times for targets ranges from approximately 0.5 to 1 min.

Endsley and Kaber (1999) computed the minimum and maximum performance in the simulation as 10 and 60 target collapses in a 60-s period, respectively. Minimum performance could result if all targets eliminated were large, and maximum performance would occur if all targets processed were small.

Targets followed one of eight approach paths from the edge of the display towards its center causing convergent-type movement. Targets could collide on the same approach path if one was traveling faster than another, or they could collide on adjacent approach paths as they neared the center of the display. This feature added to the task complexity, based on the interaction of the different target characteristics (i.e., speed and size).

In general, Multitask© is a cognitive task involving operator judgments and projections on the temporal and spatial relations of targets, as well as interpretation of target characteristics for prioritization for processing and decision making in target selection. In order to optimize performance, subjects need to develop a strategy accounting for tradeoffs among rewards associated with attending to a target and penalties associated with missing a target or allowing two targets to collide, and assessments of the time available and required for processing tasks.

The different methods for processing targets conformed to those LOAs in Endsley and Kaber's (1999) taxonomy, which were selected for assessment as part of this experiment. The operator interface and the responsibilities of the human and computer at the test LOAs were as follows:

Manual Control – The operator was required to: (a) continually monitor the display and the status of competing targets and their relevant attributes, (b) generate a strategy (processing order) for eliminating targets, (c) select targets for elimination accordingly by pointing to them with the mouse, and (d) implement their strategy (process the tasks) by continually depressing the mouse button over the selected target until it disappeared.

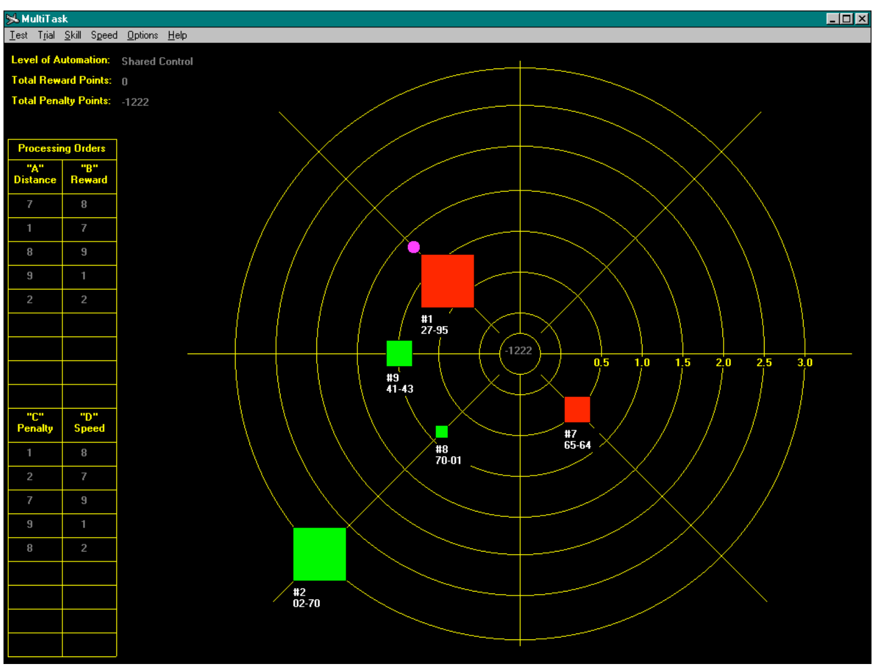

Figure 2. Multitask© simulation frozen under Shared Control.

Batch Processing – A target processing order (which was input by the operator) was also shown in the lower left corner of the display. The operator was required to: (a) generate a strategy for processing targets, and (b) select targets to be added to the processing order by depressing the numeric keys 1 through 9 on the keyboard, which corresponded to numbers tagged to the displayed targets. The computer implemented the operator's processing order by automatically collapsing each target in the list. This LOA, therefore, provided full automation of the implementation portion of the task.

Shared Control – Four processing orders were generated by the computer (based on target distance, reward, penalty and speed (see Figure 2, left side)), and were displayed to assist the operator in target selection. Additional guidance was offered by the computer in the form of a magenta dot tagged to the target (see small, filled circle adjacent to large, dark target in upper-left portion of radarscope in Figure 2) that was currently the "best" choice in terms of all variables based on an optimization algorithm that considered distance, reward, penalty and speed (i.e., [(reward – penalty) / (distance / speed)]. The operator was required to: (a) generate a processing strategy, which could be her own or could be based on the computer guidance, (b) select targets to be collapsed using the mouse, and (c) implement the strategy (process the targets) by clicking the mouse button once over the desired target to activate automated processing. This LOA provided joint human/computer generation of decision options (strategies) and joint implementation of the human decision.

Blended Decision Making – The same information, as provided under Shared Control, was presented along with a column that allowed for a processing order to be entered by the operator (as under Batch Processing). The operator and computer both generated strategies for eliminating targets (as above); however, the computer selected the processing order to be implemented. The order selected by the automation could be over-ridden by the operator at any time (if she did not agree with the computer's choice) by depressing the keys A (distance), B (reward), C (penalty), D (speed) and E (operator) corresponding to the desired order (see top cells of "Processing Orders" matrix in Figure 2). The computer implemented the selected processing order by automatically collapsing targets on the list one-at-a-time. Blended Decision Making provided a higher LOA by incorporating computer selection with human veto.

Supervisory Control – This mode offered automation of all functions with human over-ride capability. Therefore, the computer: (a) generated a processing strategy by taking into account all target variables, (b) selected targets for elimination and (c) implemented the strategy by automatically collapsing targets one-at-a-time. The operator could intervene in the control process, if she thought the computer was not efficiently eliminating targets. Operator intervention was accomplished by depressing a key, which temporarily shifted (for 1 min) the LOA to Blended Decision Making. The operator could return to automation of all functions before the end of the temporary shift in LOA by depressing a second key. This LOA was, therefore, representative of many supervisory control systems in which the system is mostly automated, but human monitoring and intervention is expected.

Full Automation – In this mode all functions comprising: (a) processing order generation, (b) target selection, (c) strategy implementation (target elimination), and (d) system monitoring were performed by the computer. Operator intervention was not permitted. Therefore, under Full Automation the operator could only observe system performance. (It is important to note that the automation was only as good as the target-processing algorithm included in the Multitask© software (see Shared Control above). In general, the system was not capable of identifying or considering potential target collisions in formulating processing plans, which was critical to ensuring all implementation efforts were worthwhile. Therefore, Full Automation allowed for good, but not perfect, performance.)

Each subject was randomly assigned to one of five groups corresponding to the automated settings of the simulation, including Batch Processing, Shared Control, Blended Decision Making, Supervisory Control and Full Automation. The LOAs were adaptively applied to the task with each level being dynamically allocated in juxtaposition with Manual Control according to different predetermined time schedules. These schedules and the extent to which each experimental trial was automated are detailed in the Procedures Section.

3.2.2. *Secondary Gauge Monitoring Task*

The gauge-monitoring task was designed based on the monitoring subtask of the MAT Battery. It presented a fixed-scale display with a moving pointer (see Figure 3) and required operators to monitor pointer movements to detect when a deviation occurred from a central "acceptable" region into peripheral "unacceptable" regions. The task was presented on a monitor separate from that used to present the Multitask© simulation and required subjects to correct for pointer deviations by depressing keys on a keyboard facilitating upward or downward motion of the pointer. This task was psychomotor in nature, involving subject monitoring, condition diagnosis and action. Performance was recorded as the ratio of the number of unacceptable pointer deviations detected to the total number of deviations (i.e., the hit-to-signal ratio).

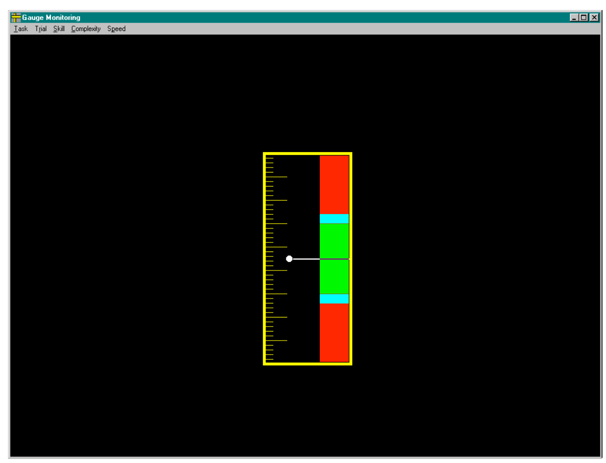

Figure 3. Gauge-monitoring task display.

The task was presented to subjects as an embedded secondary task; that is, they were instructed to maintain attention on both the Multitask© simulation and the gauge monitoring task. The task was included in the experimental scenario to provide a realistic loading of operators in that human performance with automated systems usually involves multiple tasks; thus, encouraging reliance on the automation and possibly complacency. It was also considered as an analog to many ancillary activities associated with radar monitoring in military applications (e.g., communication with other radar operators – target handoff, etc.).

3.3. *Apparatus*

The synthetic tasks were presented using two Pentium®-based PCs, two 17-in graphics monitors, two standard keyboards, and a mouse integrated with the system running the Multitask© simulation. The monitors operated at 60 Hz under 1024 × 768 resolution with refresh rates of 30 frames/s.

3.4. *Experimental Design*

A mixed between-within experimental design was used in this study. The Multitask© LOA was manipulated as a between-subjects variable and the time schedule of manual and automated control allocations was manipulated within-subjects. There were five settings of LOA, including Batch Processing, Shared Control, Blended Decision Making, Supervisory Control and Full Automation, as described in the Multitask section. There were also five DFA schedules, including one without automation; three schedules dictating “low”, “medium” and “high” automation exposure; and one completely automated schedule. Subjects experienced the various schedules of manual and automated control in random order.

3.5. *Procedures*

Subjects were initially familiarized with the procedures and equipment, including the SAGAT (Endsley, 1988) and the NASA Task-Load Index (TLX) (Hart & Staveland, 1988). As in Endsley and Kaber’s (1999) study, the SAGAT was used to measure operator SA by freezing the tasks at random points in time during the experiment, blanking all visual display screens and administering queries concerning both the current and future states of the system. The SAGAT evaluated subject perception (Level 1 SA), comprehension (Level 2 SA) and projection (Level 3 SA) regarding information displayed during the tasks by comparing subject responses to queries on each level with actual situation data recorded by the PCs running the simulations (Endsley, 1995, 1988). The percentage of correct responses was then calculated for all queries.

Subjects completed NASA-TLX demand component rankings and then were trained in Manual Control of the Multitask© and the gauge-monitoring simulation. The instructions to subjects included detailed descriptions of both the Multitask© and gauge interfaces, as well as the functionality of Multitask© under the various LOAs. The interface controls were demonstrated to subjects and they were then permitted to practice the tasks for 20 min. A 2-min rest period was provided and was followed by additional Multitask© and gauge-monitoring training for 20 min at the LOA to which a subject had been assigned. This training was followed by a 5-min break.

All subjects were required to complete five 60-min trials, three of which involved AA, one completely Manual Control trial (the DFA schedule without automation), and one requiring subjects to perform under their assigned LOA for the entire trial (the maximum automation allocation cycle time).

A model-based approach was taken to AA in this study not only to allow for examination of the effects of periodic automation on task performance, but to also assess the specific effects of the duration and frequency of AA allocations. This implementation of AA is consistent with previous work (Parasuraman, 1993; Scallen et al., 1995) and provided for better experimental control for investigating the desired issues than might have been achievable with other approaches to triggering DFAs strictly based on operator states (e.g., performance measurement, psychophysiological assessment, etc.). Research on the effects of various AA triggering strategies is also important but beyond the scope of this work. During the AA trials, subjects performed the task for predefined periods at the assigned LOA interspersed with periods of Manual Control according to one of three DFA schedules (see Figure 4). The schedules were designed to vary the distribution of automation and Manual Control during a trial and included:

(1) Low Automation Allocation Cycle Time (AACT) – involved 3 allocations of automation at the assigned LOA, interspersed with Manual Control. Automation allocations occurred at regular intervals of 16 min and lasted for 4 min. Therefore, 20% of the trial was automated.
(2) Medium AACT – also involved 3 allocations of automation at the assigned LOA with the duration of automation allocation periods set at 8 min. The interval between automation allocations was 12 min. Forty percent of the trial was automated.
(3) High AACT – three allocations of automation at the LOA to which a subject was assigned were provided. The duration of automation allocations was set at 12 min. and an automation allocation occurred every 8 min. Therefore, 60% of the trial was automated.

The Manual Control trial represented a schedule when the AACT was set to 0 min (the “None” AACT schedule), meaning no AA allocations were made. The completely automated trial represented a schedule when the AACT was equal to the task time of 60 min (the “Maximum” AACT schedule), meaning the task was automated at the assigned LOA for the entire trial period.

Subjects were informed in advance that Manual Control was to be augmented by automation allocations (on effected trials) and that they were to monitor the Multitask© display carefully to detect and respond to any and all automation allocations (as indicated by changes in dynamic interface features). However, no information concerning the frequency or durations of automation allocations within a trial was provided in order to prevent advanced preparation. All automation allocations were made salient to subjects by an audio tone and a display of the name of the LOA assigned, which was provided in a “Level of Automation” data field on the Multitask© display (see Figure 2, upper-left corner). Subjects were required to operate at the designated LOA for the scheduled duration of the AA period before returning to Manual Control, as shown in Figure 4.

In addition to the AA allocations, six task freezes were dispersed throughout each trial to administer SAGAT queries in order to assess the effect of LOA and AA on SA (also see Figure 4). Stops occurred at random points in time, with half occurring during Manual Control periods and half during automation periods. When a freeze occurred, subjects responded to an electronic form of the SAGAT queries. It included: (1) color and size identifications for each target (Level 1 SA); (2) four questions

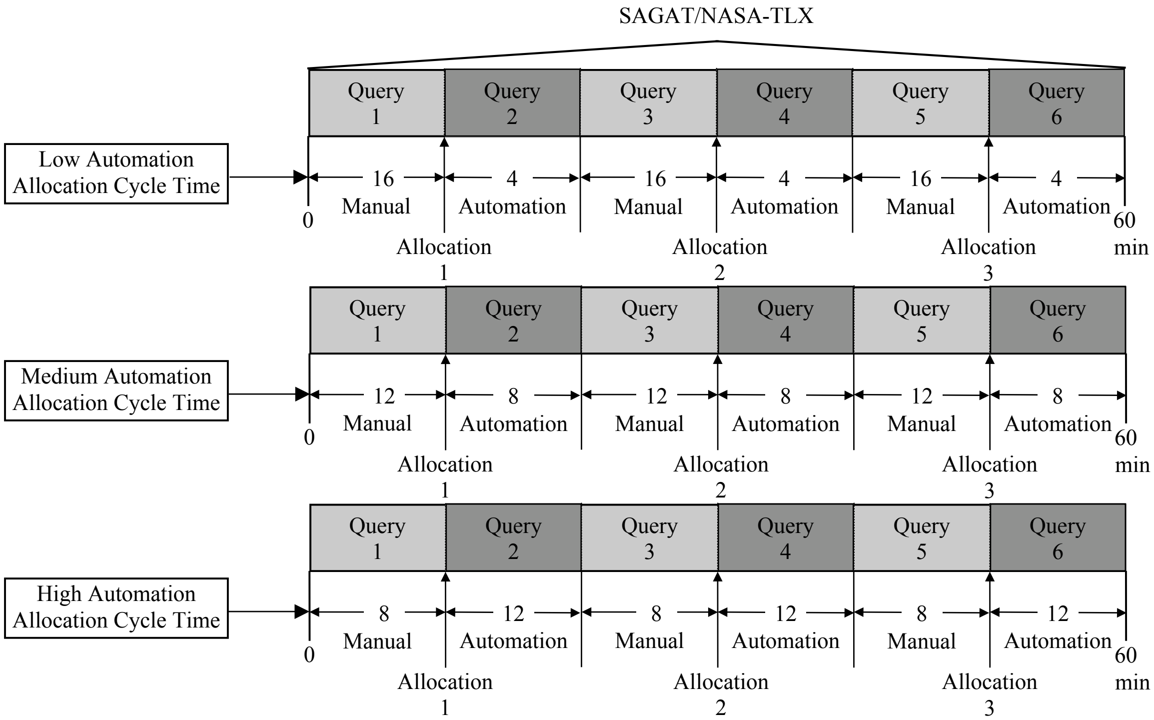

Figure 4. Schedules for AACT trials.

concerning the reward, penalty, speed and distance of targets to the deadline (Level 2 SA); and (3) one question concerning which target, of all the targets on the Multitask© display at the time of the freeze, would reach the deadline next (Level 3 SA). Each stop lasted until subjects completed the queries.

Subjects were not provided with knowledge of how many freezes were to occur during a trial or inter-freeze intervals in order to prevent advanced preparation for the quizzes. It is possible that subjects might have guessed at the number of freezes and when freezes would occur in later trials, as part of their participation. However, they could not have been certain of this, as they had no idea the number of freezes would be the same across trials. Furthermore, the timing of each freeze was randomly determined within the periods of Manual Control and automation during a trial, making prediction of the exact freeze time virtually impossible. Such freezes to collect SAGAT data (or even the possibility of such freezes) have not been found to affect subject performance (Endsley, 1995, 2000).

Directly following a SAGAT freeze, subjects completed a NASA-TLX demand component rating form, including mental, physical, temporal, performance, frustration and effort ratings. In this way, TLX scores were captured to describe workload over the course of a trial and not simply the average workload perceived at the close of a trial, as is typically the case with this measure. After a freeze, the task was resumed until trial completion. All trials were performed in an environmental chamber adjusted to normal room conditions (26 C) in order to block-out extraneous distractions (e.g., noise or interruptions). In total, subjects participated in 3 experimental sessions of 2 hr and 30 min each. The first session was used for training and a single test trial. The remaining two sessions involved two test trials with a 10-min rest period between them.

3.6. *Data Analysis*

The primary task performance measures included the number of targets collapsed, expired and collided and were recorded at 1-min intervals throughout the 60-min test periods during all 150 trials (5 LOAs × 6 subjects within LOA × 5 AACTs) yielding 9000 observations per response. (Given the average frequency/timing of target collapses, collisions and expirations in the Multitask© simulation, a shorter interval for data recording would have yielded many samples with zero observations and most likely would not have increased the accuracy or sensitivity of our analysis.) Since the target events determined the total reward and penalty points assessed, the pattern of results on, for example, rewards was identical to the pattern of results on collapses. Consequently, only analyses of the target event measures are presented here.

There were a total of 900 data points on each SAGAT query and overall workload resulting from the 6 stops as part of each trial. With respect to secondary task performance, the rate of error detection was averaged at 4-min intervals across the 150 test trials producing 2250 observations.

All data sets (primary and secondary task performance, SA and workload) were divided into two subsets for analysis, including performance during Manual Control periods and performance at each LOA allocated during the AA trials and the completely automated trial. All observations were analyzed through a two-way Analysis of Variance (ANOVA) with LOA and AACT as between- and within-subjects variables, respectively. For the Manual Control data, the LOA level listed represents the LOA that directly preceded a manual performance period (as this is relevant to SA and OOTL problems in manual performance periods immediately following).

In order to rectify violations of the underlying assumptions of the ANOVA in the data sets, transforms were applied to the various response measures according to the procedures described by Neter et al. (1990 pp. 142-146), including logarithmic transforms of all Multitask© performance measures and the NASA-TLX overall workload score, as well as arcsine transforms on the SAGAT response and secondary task performance measure. In particular, the performance measures were transformed due to non-normality of the data as revealed by a significant Shapiro-Wilks test and a non-linear trend of the sorted residuals against the expected response values in a normal probability plot. Residual plots against the levels of the various independent variables (LOA and AACT) also indicated violations of the constant variance assumption of the ANOVA. Although the results reported below are on the transformed responses, all graphs present response means at the various settings of the predictors in original units in order to promote ease of interpretation and understanding (Neter, Wasserman, & Kutner, 1990, p. 147). Finally, correlation analyses were conducted to establish any significant relationships among the various primary and secondary task measures.

4. Results

4.1. *Primary Task Manual Control Performance*

Analysis of Variance results on the performance data collected during the completely manual trial and Manual Control periods as part of the AA trials revealed no significant effects of LOA, AACT or the interaction of these variables on the log transform of the number of targets collapsed, expired or collided in the Multitask© simulation. Results of an ANOVA on (1) the arcsine transform of Level 1, 2 and 3 SA, and the rate of error detection in the secondary monitoring task; and (2) the log transform of NASA-TLX scores collected during these periods also revealed LOA, AACT and the LOA × AACT interaction to be insignificant in effect. Manual performance was, therefore, not affected by either the LOA or AA approach used in the primary task. Unlike in Endsley and Kaber's (1999) study, subjects in this experiment were instructed to view the Manual Control periods as opportunities to enhance overall system functioning. They were not to see them as automation failures. The lack of a LOA or AACT effect on manual performance may have been due to this instruction.

The remainder of this section presents results on data collected during automated performance periods. The figures presenting mean performance, SA and workload at the various LOAs (Figures 5-8) include the means for the None AACT trial and Manual Control periods during the AA trials to allow for illustrative comparison.

4.2. *Primary Task Performance at Assigned LOAs*

The findings of ANOVAs on all responses observed during subject performance under each LOA × AACT combination (p values) are summarized in Table 3. (Details on the significant F tests, including degrees of freedom and test statistics, are included in the text as part of this subsection and the subsections on SA, workload and secondary-task performance results.) The Table reveals significant main effects and the presence of interactions across the majority of responses including target collapses, expirations and collisions, Level 2 SA, NASA-TLX scores and the hit-to-signal ratio for the gauge-monitoring task.

Results of ANOVAs on the log transform of target collapses ($F(4,25) = 31.11$, $p < 0.0001$), expirations ($F(4,25) = 8.37$, $p < 0.0002$) and collisions ($F(4,25) = 18.22$, $p < 0.0001$) revealed a significant main effect of LOA. The AACT was also significant in terms of the number of targets processed ($F(3,25) = 28.1$, $p < 0.0001$). An interaction effect was present for the log

Table 3. Summary of results on analysis of automation performance, SA and workload data.

Response Measure	Predictor Variable					
	AACT		LOA		LOA × AACT	
Target Collapses	$p = 0.0001$	**	$p = 0.0001$	**	$p = 0.0001$	**
Target Expirations	$p = 0.5846$		$p = 0.0002$	**	$p = 0.0553$	
Target Collisions	$p = 0.2741$		$p = 0.0001$	**	$p = 0.0001$	**
Level 1 SA	$p = 0.4541$		$p = 0.2648$		$p = 0.8665$	
Level 2 SA	$p = 0.4268$		$p = 0.0021$	**	$p = 0.0331$	*
Level 3 SA	$p = 0.7018$		$p = 0.3387$		$p = 0.1468$	
NASA-TLX	$p = 0.0011$	**	$p = 0.1971$		$p = 0.0469$	*
Rate of Error Detection	$p = 0.0181$	*	$p = 0.2928$		$p = 0.0003$	**

* - Significant at the $\alpha = 0.05$ level.

** - Significant at the $\alpha = 0.01$ level.

transform of collapses ($F(12,75) = 26.98$, $p < 0.0001$) and collisions ($F(12,75) = 5$, $p < 0.0001$).

The mean number of targets collapsed increased linearly with increasing automation cycle time. Duncan's Multiple Range (MR) test revealed significant differences ($p < 0.05$) among the Low, Medium and High AACT settings, but not between the High and Maximum AACT. The shortest AACT yielded the worst performance. These findings can be attributed to the computer assistance provided to operators across the various LOAs during automation allocations.

Low and intermediate automation (Batch Processing (LOA 3) and Blended Decision Making (LOA 6)) produced, on average, more collapses and reduced the mean number of targets expirations and collisions, respectively, as compared to high-level automation (Full Automation) and Manual Control. In general, target expirations appeared to vary as a "U" function of LOA with Manual Control and Full Automation producing, on average, worse performance. Duncan's tests on the primary performance measures revealed each LOA to differ significantly ($p < 0.05$) from every other in terms of mean target collapses; all LOAs to differ significantly ($p < 0.05$) in terms of expirations, save Shared (LOA 4) and Supervisory Control (LOA 9); and collisions to differ significantly ($p < 0.05$) between Batch Processing (LOA 3) and all other levels.

The mean numbers of target collapses across subjects, as a function of the LOA × AACT interaction, are shown in Figure 5. Under all cycle times, collapses varied as a function of LOA, with peak numbers at Batch Processing (LOA 3) and Full Automation (LOA 10). The mean number of target collisions, like collapses, also varied with LOA revealing slight peaks at Manual Control (LOA 1) and Batch Processing (LOA 3). Duncan's MR test was conducted on the log transform of both responses for each LOA × AACT combination. In general, Batch Processing (LOA 3) under the High AACT trials produced superior performance ($p < 0.05$) compared to all other conditions. It was possible for operators assigned to the Batch Processing condition to outperform the system operating under Full Automation because the automation algorithm did not consider potential or imminent target collisions in planning a processing schedule. Therefore, the computer might have selected and begun to process a target that ultimately collided with another target. Since Multitask© only permitted processing of one target at a time, this type of automated selection essentially represented wasted implementation time. The Batch Processing level also produced good performance during Medium AACT trials and the fully automated trial, but performance was never as good as in the AA trials. Duncan's test also revealed Shared Control (LOA 4) during the fully automated and Low AACT trials to produce significantly fewer ($p < 0.05$) target collapses than all other LOA × AACT combinations.

With respect to target collisions, Duncan's tests indicated that Batch Processing (LOA 3) as part of the High AACT schedule produced the worst performance. Collision prevention was also poor under Batch Processing (LOA 3) as part of the other AA trials and the fully automated trial. Next to these conditions, Supervisory Control (LOA 9) as part of the Medium AACT schedule produced a significantly higher, average number of collisions ($p < 0.05$) than all other LOA × AACT combinations. Otherwise, the mean number of target collisions was essentially constant across Shared Control (LOA 4), Blended Decision Making (LOA 6) and Full Automation (LOA 10) as part of various AA trials and the fully automated trial. Blended

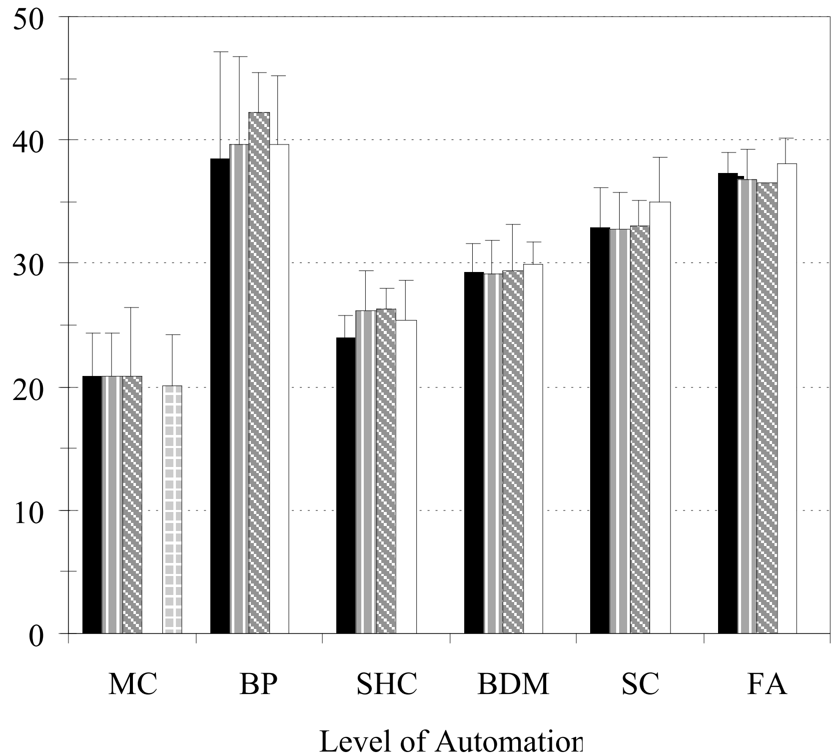

Figure 5. Plot of mean number of target collapses for each LOA × AACT combination. (Note: MC = Manual Control; BP = Batch Processing; SHC = Shared Control; BDM =Blended Decision Making; SC = Supervisory Control; and FA = Full Automation. Error bars represent ±1-standard deviation.)

Decision Making (LOA 6) as part of the Low AACT schedule produced the best performance in terms of preventing target collisions.

4.3. *Primary Task Situation Awareness*

Results of an ANOVA on a 5 × 4 (LOA × AACT) model of the arcsine transform of Level 1, 2 and 3 SA indicated that LOA was significant in effect ($F(4,25) = 5.68$, $p = 0.0021$) on the average percent correct responses to task comprehension queries (Level 2 SA). The LOA × AACT interaction was also present ($F(12,75) = 2.03$, $p = 0.0331$) for Level 2 SA. The percent correct responses to queries varied with LOA across AACT and did not fall below chance, which was 20% for each query, in any of the conditions.

Graphical analysis revealed subject comprehension of target characteristics in relation to task goals (maximizing reward points and minimizing penalties) to peak at an intermediate LOA (Shared Control (LOA 4) and Full Automation (LOA 10)), while low and high-level automation (Batch Processing (LOA 3) and Supervisory Control (LOA 9)) produced substantially lower SA. Duncan's test confirmed these observations and indicated Batch Processing and Supervisory Control to significantly differ ($p < 0.05$) from all other levels.

Figure 6 shows the mean percent correct responses to Level 2 SA queries across subjects, as a function of LOA × AACT interaction. Operator task comprehension appeared to peak under Shared Control (LOA 4) as part of the Low and Medium

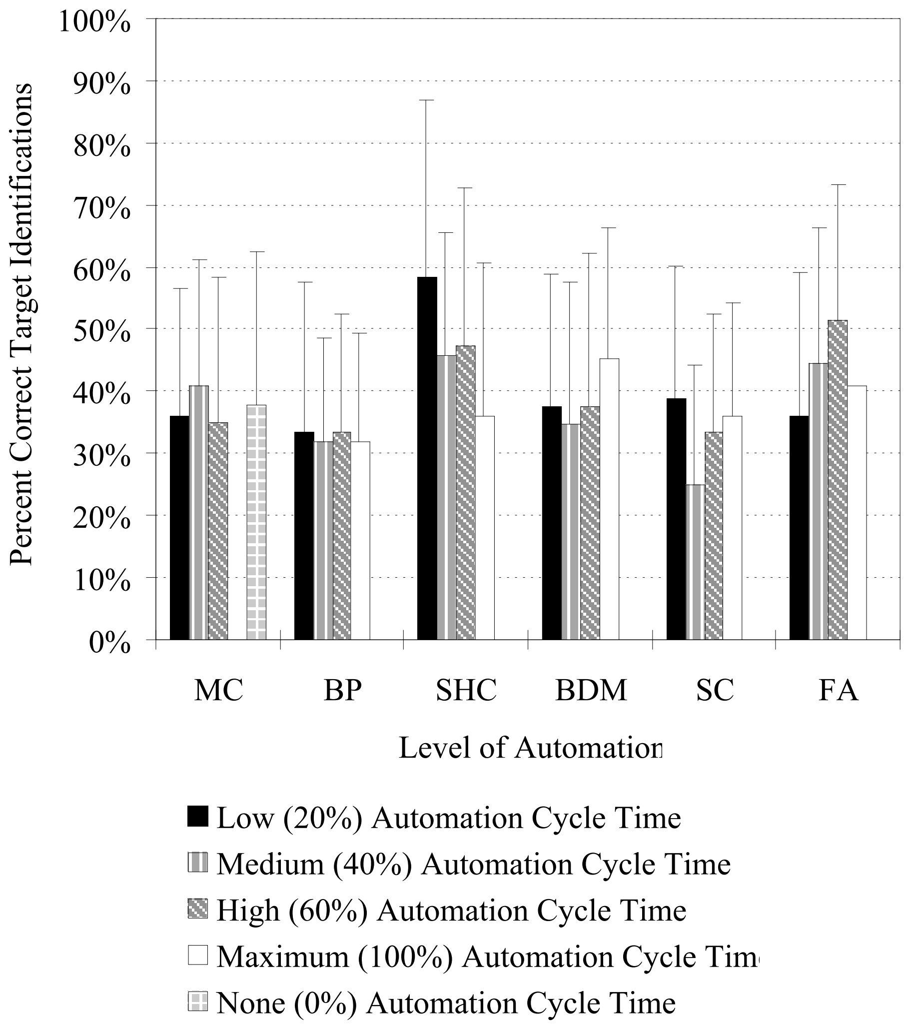

Figure 6. Plot of mean percent correct responses to Level 2 SA queries for each LOA × AACT combination. (Note: MC = Manual Control; BP = Batch Processing; SHC = Shared Control; BDM =Blended Decision Making; SC = Supervisory Control; and FA = Full Automation. Error bars represent +1-standard deviation.)

AACT schedules. When the High automation cycle time setting was used, Full Automation (LOA 10) produced substantially higher SA. It can also be noted from the graph that Manual Control (LOA 1) as part of the completely manual trials produced higher SA than Manual Control during the Low and Medium AACT trials, but lower SA than during the High AACT trial. On average, SA during completely Manual Control was never as good as the best SA during Manual Control periods as part of AA trials (High AACT). Duncan's MR test indicated Shared Control (LOA 4) at the Low-cycle time and Full Automation (LOA 10) at the High-cycle time to produce the best SA. The mean percent correct responses to queries were significantly lower ($p < 0.05$) under Batch Processing (LOA 3) and Supervisory Control (LOA 9) during the AA trials (Medium and High AACT schedules) than the average for all other LOA × AACT combinations.

4.4. *Ratings of Primary Task Workload*

Results of an ANOVA on the NASA-TLX response indicated a significant main effect of AACT ($F(3,25) = 5.92$, $p = 0.0011$) and a significant interaction of LOA × AACT ($F(12,75) = 1.91$, $p = 0.0469$). Not surprisingly, NASA-TLX scores tended to decrease with increasing duration of automation. Duncan's tests on the logarithmic transformed TLX scores revealed significant differences ($p < 0.05$) among all AACT settings with peak workload occurring under the Low-cycle time and the minimum mean score at the Maximum AACT.

The mean NASA-TLX scores across subjects, as a function of AACT and LOAs are shown in Figure 7. In general, workload ratings appeared to peak under Shared Control (LOA 4) during the AA and fully automated trials. Manual Control (LOA 1) performance during the completely manual trials yielded, on average, lower TLX scores, than Manual Control (LOA 1) during the AA trials, specifically when the Low and High AACT schedules were administered. Duncan's MR test revealed that Shared Control (LOA 4), in general, produced the greatest perceived workload in comparison to all other LOA × AACT conditions. Batch Processing (LOA 3) as part of the Low AACT schedule also produced significantly higher ($p < 0.05$) ratings of workload than all other experimental conditions, save those involving Shared Control (LOA 4). As one might expect, the lowest workload ratings were observed for Full Automation (LOA 10) at the Maximum-cycle time, according to Duncan's tests.

4.5. *Embedded Secondary Task Performance*

Results of an ANOVA on the 5 × 4 (LOA × AACT) model of the arcsine transform of operator performance in the secondary (gauge monitoring) task indicated AACT ($F(3,25) = 3.56$, $p = 0.0181$) and the LOA × AACT ($F(12,75) = 3.59$, $p = 0.0003$) interaction to be significant in effect. Not surprisingly, error detection rates increased with increasing durations of automation in the primary task. According to Duncan's test, the Maximum (100%) automation cycle time of the primary task produced the greatest number of error detections in the secondary task, while the fewest errors were detected during the Low AACT schedule and completely manual trials. The Medium and High AACT schedules were not significantly different ($p > 0.05$).

Figure 8 shows the mean rate of error detection in the secondary task as a function of LOA of the primary task and AACT. From the graph, operator performance in gauge monitoring appeared to peak when the Multitask© simulation was automated under

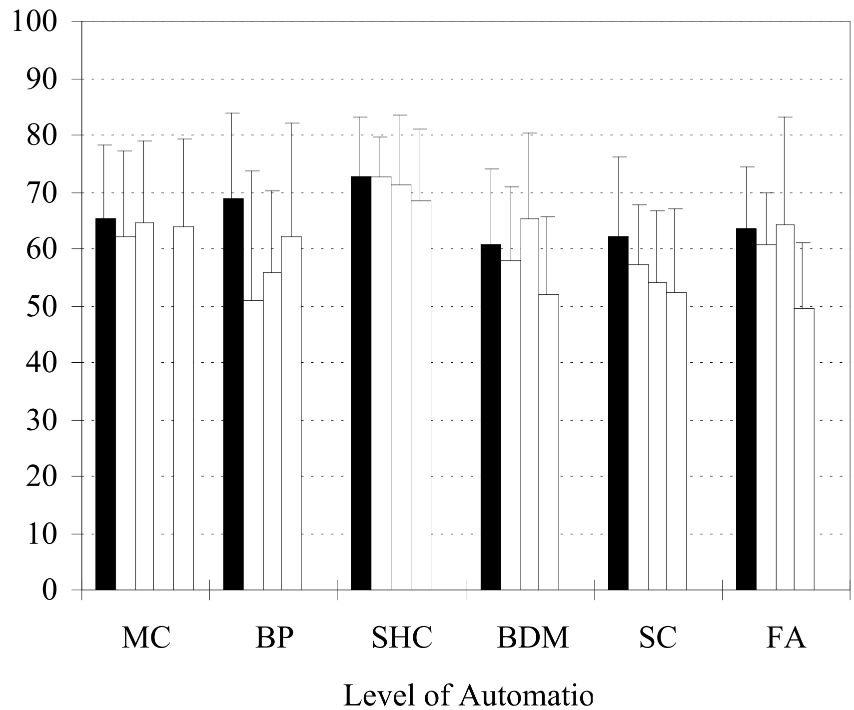

Figure 7. Plot of mean NASA-TLX overall workload for each LOA × AACT combination. (Note: MC = Manual Control; BP = Batch Processing; SHC = Shared Control; BDM =Blended Decision Making; SC = Supervisory Control; and FA = Full Automation. Error bars represent +1-standard deviation.)

high-level automation (Supervisory Control (LOA 9)) for Medium- and High-cycle times. Duncan's MR tests revealed that Full Automation (LOA 10) of the primary task as part of the Maximum AACT schedule caused the best performance in the secondary task followed by the use of Batch Processing (LOA 3) during the Medium AACT trial. During Manual Control (LOA 1) of the Multitask© simulation, performance in gauge monitoring was, on average, greatest when the High AACT schedule was followed. This seems logical because of all the trial schedules involving Manual Control (LOA 1), the High AACT schedule provided computer assistance for the greatest percentage of time on task. Consequently, operator perceptual resources required by the Multitask© may have been freed-up more often for detecting errors in the gauge task. Manual Control (LOA 1) during the completely manual trials appeared to produce a lower average error detection rate than Manual Control (LOA 1) as part of the AA trials. Duncan's test also revealed Shared Control (LOA 4) and Blended Decision Making (LOA 6) as part of the Low and Medium AACT schedules to produce significantly lower ($p < 0.05$) mean rates of error detection in the gauge monitoring task than all other LOA × AACT combinations, save Batch Processing (LOA 3) as part of the Low and High AACT schedules.

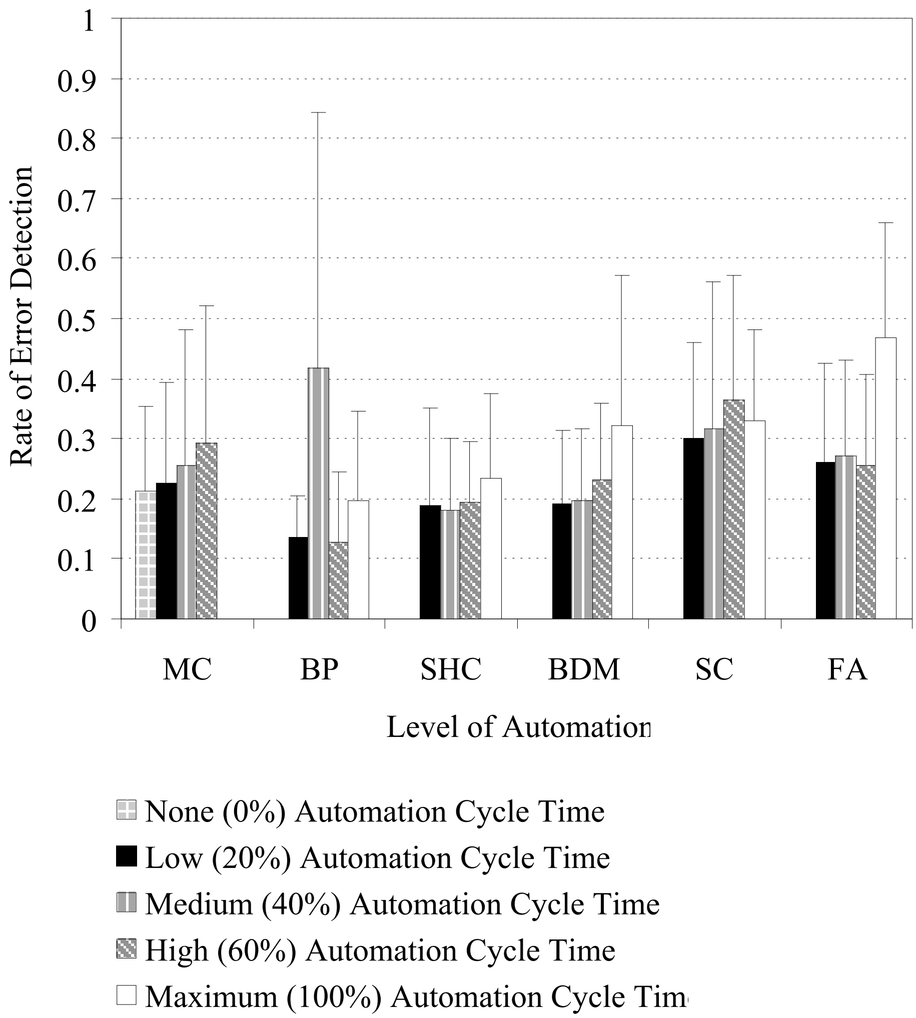

Figure 8. Plot of mean rate of error detection for each LOA × AACT combination. (Note: MC = Manual Control; BP = Batch Processing; SHC = Shared Control; BDM =Blended Decision Making; SC = Supervisory Control; and FA = Full Automation. Error bars represent ±1-standard deviation.)

4.6. *Response Measure Correlation Analyses*

Pearson-product moment coefficients were determined for all pairs of response measures observed during the study to determine: (1) whether workload reductions were accompanied by improvements in SA; (2) whether there were tradeoffs among the two tasks in terms of performance, SA or workload; and (3) whether SA on the primary task may have been associated with poor performance in the secondary task.

Significant relationships of interest to this research occurred among Level 2 SA queries and NASA-TLX scores (r = 0.4175, p = 0.0424), and NASA-TLX scores and the rate of error detection in the secondary task (r = -0.6993, p = 0.0001). As workload increased, operator task understanding appeared to improve. This may have been due to greater subject involvement in the system control loop at lower LOAs (e.g., Shared Control (LOA 4)), or due to Manual Control (LOA 1) allocations during functioning at high LOAs (e.g., Full Automation (LOA 10)) as part of AA trials. As perceptions of primary task workload increased, operator performance in gauge monitoring decreased.

There were also significant correlations between Level 2 SA and primary task performance (the number of target collisions) (r = 0.4518, p = 0.0267); however, performance in the secondary task was only weakly correlated with Level 2 SA (operator comprehension of the states of the primary task). The number of target collisions permitted by an operator decreased as their understanding of target information improved. Increased levels of operator SA may have improved performance.

5. Discussion

5.1. *Primary-Task Performance*

In general, the pattern of performance in the Multitask© simulation under the various LOAs replicated the results obtained by Endsley and Kaber (1999), even though different AA strategies were employed. That is, LOA appeared to be the driving factor in human-computer performance in the dynamic control task; whereas, automation cycle time had comparatively little effect, as can be seen in the interaction plot on target collapses. With reference to the LOA taxonomy in Table 2, improvements in performance in terms of the number of targets addressed occurred at LOAs involving human strategizing (generating and selecting target processing plans) with computer aiding in the implementation aspect of the task (Batch Processing (LOA 3)), as compared to all other levels that added automation to the roles of option generation and selection (Shared Control (LOA 4) and Blended Decision-Making (LOA 6)).

As we hypothesized, when low-level automation, Batch Processing (LOA 3), was combined with a greater percentage of time-on-task being automated (the High AACT condition), it produced better performance. We expected the longer automation cycles as part of the AA condition to improve performance. Batch Processing (LOA 3) allowed for operator advanced queuing of targets for processing by the computer versus Blended Decision Making (LOA 6), which also permitted operator entry of a target processing order. However, the computer had a role in generating strategies as well, which may have distracted operators from task performance. The performance results for Shared Control (LOA 4), identified this intermediate level as the worst among all conditions in terms of target processing. This may reveal the interaction of the short-cycle automation (the Low

AACT condition) with the specific characteristics of the LOA as being problematic. The short control cycles as part of some AA trials were expected to have a negative impact on performance.

Although Batch Processing (LOA 3) may have been effective for eliminating targets, it also yielded the greatest number of collisions as part of the High AACT schedule. During advanced queuing of targets for computer processing, operators may have focused their attention on the goal of maximizing rewards (processing targets) versus preventing penalties associated with task conflicts. With the advanced queuing capability, operators may not have been operating in the "moment" and missed critical events because of the automation. It is possible that automation of certain system information processing functions may cause operators to be essentially OOTL with respect to certain task processes, while focused on others.

In agreement with one of our initial hypotheses, compared to Batch Processing (LOA 3) performance improvements in the form of decreases in the number of target collisions occurred at intermediate LOAs requiring joint human-computer generation of processing plans (Shared Control (LOA 4) and Blended Decision-Making (LOA 6)), specifically during the AA trials with Low and Medium AACT. Although these LOAs did not support task processing as well as those involving purely human strategy generation, the computer decision guidance may have caused operators to behave more conservatively and pay closer attention to penalties associated with disregarding tasks.

5.2. *Situation Awareness*

In general, as we hypothesized situation understanding improved under intermediate level automation (Shared Control (LOA 4)), however, contrary to our expectation based on Endsley & Kiris' (1995) findings, comprehension also improved under Full Automation (LOA 10) but this was during the AA trials involving periods of Manual Control. Situation awareness in trials only involving Full Automation (the Maximum AACT condition) was significantly worse than in the AA trials. In general, aspects of our results were similar to both Endsley and Kiris' (1995) and Endsley and Kaber's (1999) findings. Full Automation (LOA 10) removed subjects from the Multitask© control loop, however, the reduced taskload may have freed-up cognitive resources for information perception and integration in working memory while keeping task goals in mind (resulting in higher SA under full automation). Periods of Manual Control during the AA trials reacquainted operators with the current state of the system and increased SA in comparison to completely automated performance.

Shared Control (LOA 4) maintained operator task involvement by requiring option selection of the operator and collaboration with automation in system monitoring, options generation and implementation. Even higher perceptions of workload than those observed may have been offset by providing computer assistance in these roles.

Batch Processing (LOA 3) and Supervisory Control (LOA 9) (low and high LOAs) caused the worst operator SA of target priorities and completion status across the various AACT conditions. Endsley and Kaber (1999) obtained the same result for Batch Processing (LOA 3) using the Multitask© simulation. Operator involvement in advanced cueing of targets for processing may have distracted from their comprehension of current system events, such as target collisions (task conflicts) and expirations (tasks disregarded). As previously discussed, operators may have been OOTL in terms of the implementation aspect of the task and this could have undermined their task comprehension. As hypothesized, low SA under Supervisory Control (LOA 9) and the High AACT condition, in particular, may be attributed to operator OOTL performance under this condition combined with stress due to concern with when intervention may be necessary to optimize performance (e.g., prevent target collisions undetected by the automation). There did not appear to be evidence of short-cycle automation as part of AA trials negatively affecting SA as we speculated; however, this effect was observed with the performance measures.

As stated, the highest Level 2 SA occurred under Full Automation (LOA 10) as part of High-cycle time automation. Interestingly, this LOA × AACT combination is equivalent to that investigated by Parasuraman (1993), who demonstrated improvements in human performance of the monitoring aspect of the MAT Battery, as compared to fully automated monitoring. This suggests that AA involving full automation and manual control may yield improved performance with high SA.

Finally, the results on SA during the completely manual trials versus Manual Control periods as part of AA indicate that some AA may always be better than none for ensuring operator comprehension of current system states. The workload reductions attributable to the AA allocations may be critical to freeing-up operator cognitive resources for task concentration and achievement of higher levels of SA. These results are comparable to the recent findings of Clamann et al. (2002), who also observed that some AA may always be better than none in terms of performance during manual control periods when AA is applied exclusively to specific human-machine system information processing functions.

5.3. *Workload*

The automation cycle time appeared to be the driving factor in changes in workload, whereas, the LOA had comparatively little effect, as can be seen in the interaction plot of NASA-TLX scores. As expected, the long cycle automation as part of the AA conditions yielded lower operator workload. Interestingly, for the low automation condition of Batch Processing (LOA 3), subjects appeared to perceive advanced cueing of targets during an entire trial (the Maximum AACT schedule) to be more taxing than periodic Batch Processing (LOA 3) combined with manual control during the AA trials. That is, full-time monitoring of automation involved more workload. This suggests that forms of automation that allow operators to work ahead of actual system processing over extended time periods may ultimately be perceived as posing greater operator workload than adaptive systems providing the same capabilities and requiring operators to track function allocations. Examples of batch processing operations include job schedulers in manufacturing systems and pilots operating with flight managements systems.

5.4. *Secondary-Task Performance*

In general, secondary monitoring performance improved with an increasing duration of automation in the primary task; however, the Multitask© LOA was not a driving factor in secondary task performance. The secondary task measure essentially validated the results on the NASA-TLX scores. This was not surprising, as substantial research has demonstrated the validity and reliability of secondary task measures of workload (see Wickens (1992, pp. 393-396)), and was supported by the correlation analysis revealing a highly significant relationship among NASA-TLX scores and secondary-task performance.

6. General Discussion and Conclusions

This work was intended to expand the current understanding of LOA and AA as approaches to human-centered automation with the objective of optimizing human-machine system performance. It investigated the interaction of these approaches in automating a dynamic control task and the implications for operator performance, SA and workload. The experiment represents basic cognitive engineering research aimed at providing insights into the use of dynamic control allocations across a broad range of LOAs during complex control task performance for the purposes of facilitating SA and managing operator workload. The experiment was conducted in a controlled laboratory setting using abstract simulations in order to develop general results potentially applicable to a broad range of domains. The work was not intended to generate results, or design guidelines, for a specific domain, such as aircraft piloting or ATC, and, therefore, care should be taken in making applied inferences on the basis of the findings. Further research is needed to explore the generalizability of these findings to more realistic tasks.

In general, this research is supportive of the use of LOA and AA approaches to facilitate human-automation interaction and to promote operator performance and SA through meaningful involvement in systems control. In a dual-task scenario, the LOA approach appears to have a greater influence than AA on primary-task performance and operator SA, with little effect on perceived workload and secondary-task performance. Interestingly, primary task performance was more closely associated with SA than with workload.

The lack of an effect of LOA on secondary-task performance is likely attributed to the fact that across the LOAs investigated here, operators always maintained a role in the dynamic control task, whether it involved options generation, selection and implementation or monitoring under full automation. Monitoring automation has been found to involve considerable workload (Becker, Warm, & Dember, 1991; Dittmar, Warm, Dember & Ricks, 1993).

The AA approach, or the definition of a schedule of DFAs, appears to drive changes in subjective workload and secondary task performance with little effect on primary task performance and SA. Not surprisingly, when a greater percentage of primary task time was automated, greater operator perceptual resources were freed-up for observing secondary task events.

In general, the combined effect of the LOA and AA approaches on Multitask© and secondary task performance, operator SA and perceived workload was not additive in nature. The LOA yielding the best overall performance (Batch Processing (LOA 3)) did not do so at the AACT producing superior functioning (Maximum automation cycle time – i.e., always automated at the LOA). The best combination of LOA and AACT involved human strategizing combined with computer implementation (Batch Processing (LOA 3)) during high automation cycle times (12-min on cycle and 8-min off cycle). This indicates that in dynamic, multi-task environments some human manual performance is useful to overall system functioning. It was better than fully automated performance, which was also considered in this research.

The combination of intermediate LOAs (Shared Control (LOA 4) and Blended Decision Making (LOA 6)) with low and medium automation cycle times (4-min on cycle and 16-min off cycle, and 8-min on cycle and 12-min off cycle) tended to produce poor performance. With respect to previous research (Parasuraman, 1993; Hilburn, Molloy, Wong & Parasuraman, 1993; Scallen, Hancock, & Duley, 1995) on the effect automation and manual control cycle durations during AA of complex tasks, it appears that automation cycles longer than 10-min separated by manual control periods of a shorter duration may benefit overall system performance.

Although, as hypothesized, fairly low-level automation (Batch Processing (LOA 3)) with a high automation cycle time (12-min on-cycle and 8-min off-cycle) produced superior human-machine system performance, this combination was associated with poorer SA (operator task understanding) and moderate workload. (This problem was previously hypothesized to reflect problems with advanced queuing of tasks at this LOA.) Secondary task performance was also worse under these conditions than for all other LOA and AACT combinations.

On the basis of the literature review (Endsley & Kaber, 1999; Endsley & Kiris, 1995; Kaber & Riley, 1999), the combinations of LOA and AACT expected to yield improved SA and low workload in conjunction with better performance included intermediate LOAs (e.g., Shared Control (LOA 4) and Blended Decision Making (LOA 6)) combined with low and medium automation cycle times (4-min on cycle and 16-min off cycle, and 8-min on cycle and 12-min off cycle). These combinations were associated with higher SA, but produced poor (primary and secondary task) performance without reductions in workload.

In general, the results on SA and workload presented here are consistent with prior research by Endsley and Kiris (1995), who found better SA at intermediate LOAs and poorer SA at full automation with no differences in perceived workload across levels. They are also in agreement with the results obtained by Kaber et al. (2000) revealing higher SA under low to intermediate automation and lower SA at high LOAs. However, Kaber et al. (2000) did find significant differences in workload across LOAs as found here, with lower levels occurring under higher LOAs.

This investigation has demonstrated higher SA at intermediate LOAs, involving computer aiding in all task roles (options generation and implementation, and system(s) monitoring) except strategy selection (decision making) with no significant effect of LOA on operator' perceptions of workload. However, improved SA at full automation was also observed in this study. This is consistent with Endsley and Kaber (1999) who found higher SA under high-level automation. Endsley and Kaber (1999) attributed this to a relatively short trial duration and the absence of a secondary task in their experimental scenario. Neither of these two factors occurred in the present experiment, however, indicating the result may be reliable.

Although fully automated functioning essentially removes operators from the control loop with the potential of jeopardizing SA, reduced task requirements, in the form of fewer different roles to maintain, may free-up operator cognitive resources for processing information in the context of their goals. Fewer task requirements does not necessarily dictate a reduction in perceived workload, however, which may remain high during high-level automation due to operator system monitoring. This explanation is, however, counter to the findings of Kaber et al. (2000) on high-level automation. It may be possible that although their subjects experienced lower workload under full automation, they may not have been attentive to the task and experienced complacency and vigilance decrements during long-duration trials (1 hr of telerobot control).

Contrary to our expectation, performance under intermediary LOAs dictating human-computer collaboration on task functions, such as option generation and selection, appears to be generally poor across various automation cycle times. Operators tend to become distracted from task performance by forms of computer guidance or because they need to attend to additional input. Furthermore, this guidance may cause them to become doubtful of their own strategy, even when it is better.

These events seem to have an impact on secondary task performance, as well. That is, even when a high LOA is used for primary task performance, it may not reduce workload sufficiently to improve secondary task performance due to the monitoring load imposed under passive conditions.

These conclusions further refine theoretical hypotheses on the usefulness of LOA and AA approaches to human-centered automation in terms of recommendations for promoting overall system performance, operator SA and managing workload, and may be used to expand the theory of human-centered automation. Although LOA and AA were found to significantly interact in effecting human-machine system performance, operator SA, and workload, the allocation of functions to humans and automated systems (i.e., the LOA) appears to be far more important to performance and SA than the amount of time that is spent on a task under automated versus manual control (AA). Whereas, the schedule of DFAs provided during complex system operations was confirmed by this study to be a key approach to managing operator workload, in comparison to manipulating the function allocation scheme alone.

6.1. *Future Research*

Based on the results of this study and the literature reviewed, a need exists for further empirical assessment of LOA and AA strategy effects on human/machine system performance in real-world dynamic control tasks. All experimental investigations of AA (e.g., Glenn, Barba, Wherry, Morrison, Hitchcock, & Gluckman, 1994; Gluckman, Carmody, Morrison, & Hitchcock, 1993; Harris, Goernert, Hancock, & Arthur, 1994; Hilburn, Molloy, Wong, & Parasuraman, 1993; Parasuraman, 1993; Parasuraman, Bahri, Deaton, Morrison, & Barnes, 1992; Parasuraman, Mouloua, Molloy, & Hilburn, 1993; Parasuraman, Mouloua, & Molloy, 1996; Hilburn, Jorna, Byrne, & Parasuraman, 1997; Kaber & Riley, 1999; Moray, Inagaki, & Itoh, 2000) and LOA (e.g., Billings, 1991; Endsley & Kaber, 1999; Endsley & Kiris, 1995; Kaber, Endsley, & Onal, 2000; Lorenz et al., 2001; Moray et al., 2000) conducted thus far have been performed in the context of simulations.

Adaptive automation, as investigated in this study, was applied using predefined schedules of manual control allocations alternated with automated functioning. The durations of the automation allocations and the inter-allocation-intervals were determined in advance of actual operator performance to provide experimental control. A need exists for empirical examination of both a broader range and finer division of control allocation cycle times to determine whether extremely long-cycle AA produces performance benefits.

Finally, alternate approaches to DFA have been put forth including allocations of automation based on dynamic evaluation of operator workload and performance. These and other methods for applying AA based on operator physiologic state, as indicated by heart rate, or mental arousal level, as indicated by EEG signals (e.g., P300) (Pope, Comstock, Bartolome, Bogart, Burdette, 1994; Freeman, Mikulka, Scerbo, Prinzel, & Clouatre, 1999; Freeman, Mikulka, Prinzel, & Scerbo, 2000), need to be systematically compared in terms of the capability to manage operator workload and facilitate SA.

Future research needs to further explore the complex interactions and tradeoffs associated with these two approaches to better integrating humans and automation. A wide range of possibilities exist in this design space, which has to date, only been explored at the edges. As more systems become automated, in a wide variety of domains, better design guidance is needed as to the effects of AA and LOA decisions. From this research, it appears that each approach offers certain benefits. The extension of this work to a variety of realistic domains is needed, so that the robustness of these findings can be further established.

Acknowledgment

Work on this paper was supported by the National Aeronautics and Space Administration (NASA), Langley Research Center (LaRC), Crew/Vehicle Integration Branch, through Grant No. NAG1-01-002, and the NASA LaRC, Psychological/Physiological Stressors & Factors Program, through Grant No. NAG-1-0139. The grant monitors included Anna Trujillo and Lance Prinzel. The views and conclusions presented in this paper are those of the authors and do not necessarily reflect the views of NASA.

References

Becker, A. B., Warm, J. S., & Dember, W. N. (1991). Effects of feedback on perceived workload in vigilance performance. *Paper presented at the Human Factors Society 35th Annual Meeting*. Santa Monica, CA: Human Factors Society.

Billings, C. E. (1988). Toward human centered automation. In S. D. Norman & H. W. Orlady (Eds.), *Flight deck automation: Promises and realities* (pp. 167-190). Moffet Field, CA: NASA-Ames Research Center.

Billings, C. E. (1991). *Human-centered aircraft automation: A concept and guidelines* (NASA Tech. Memo. No. 103885). Moffet Field, CA: NASA-Ames Research Center.

Billings, C. E. (1997). *Aviation Automation: The Search for a Human-Centered Approach*. Mahwah, NJ: Lawrence Erlbaum Assoc.

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.

Byrne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology*, *42*, 249-268.

Carmody, M. A., & Gluckman, J. P. (1993). Task specific effects of automation and automation failure on performance, workload and situational awareness. In R. S. Jensen & D. Neumeister (eds.), *Proceedings of the 7th International Symposium on Aviation Psychology* (pp. 167-171). Columbus, OH: Department of Aviation, The Ohio State University.

Clamann, M. P., Wright, M. C., & Kaber, D. B. (2002). Comparison of performance effects of adaptive automation applied to various stages of human-machine system information processing. In *Proceedings of the Human Factors Society 46th Annual Meeting* (pp. 342-346). Santa Monica, CA: Human Factors and Ergonomics Society.

Corso, G. M., & Moloney, M. M. (1996). Human performance, dynamic function allocation and transfer of training. In R. Koubek and W. Karwowski (eds.), *Manufacturing agility and hybrid automation - I* (pp. 636-639). Amsterdam: Elsevier.

Dittmar, M. L., Warm, J. S., Dember, W. N., & Ricks, D. F. (1993). Sex differences in vigilance performance and perceived workload. *The Journal of General Psychology*, *120*(3), 309-322.

Draper, J. V. (1995). Teleoperators for advanced manufacturing: Applications and human factors challenges. *The Int. J. of Human Factors and Manufacturing*, *5*(1), 53-85.

Endsley, M. R. (1987). The application of human factors to the development of expert systems for advanced cockpits. In *Proceedings of the Human Factors Society 31st Annual Meeting* (pp. 1388-1392). Santa Monica, CA: Human Factors and Ergonomics Society.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97-101). Santa Monica, CA: Human Factors and Ergonomics Society.

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors*, *37*(1), 65-84.

Endsley, M. R. (1996). Automation and situation awareness. In R. Parasuraman & M. Mouloua (eds.), *Automation and human performance: Theory and applications* (pp. 163-181). Mahwah, NJ: Lawrence Erlbaum Associates.

Endsley, M. R. (2000). Direct measurement of situation awareness: validity and use of SAGAT. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement*. Mahwah, NJ: LEA.

Endsley, M. R., & Kaber, D. B. (1997). The use of level of automation as a means of alleviating out-of-the-loop performance problems: a taxonomy and empirical analysis. In P. Seppala, T. Luopajarvi, C. H. Nygard, & M. Mattila (eds.), *13th Triennial Congress of the International Ergonomics Association* (Vol. 1, pp. 168-170). Helsinki: Finnish Institute of Occupational Health.

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.

Endsley, M. R. & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, *37*(2), 381-394.

Freeman, F. G., Mikulka, P. J., Scerbo, M. W., Prinzel, L. J., & Clouatre, K. (1999). Evaluation of a psychophysiologically controlled adaptive automation system, using performance on a tracking task. *Applied Psychophysiology and Biofeedback*, *25*(2), 103-115.

Freeman, F. G., Mikulka, P. J., Prinzel, L. J., & Scerbo, M. W. (2000). Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology*, *50*, 61-76.

Glenn, F., Barba, C., Wherry, R. J., Morrison, J., Hitchcock, E., & Gluckman, J. P. (1994). Adaptive automation effects on flight management task performance. In M. Mouloua & R. Parasuraman (eds.), *Human performance in automated systems: Current research and trends* (pp. 28-32). Hillsdale, NJ: Lawrence Erlbaum.

Gluckman, J. P., Carmody, M. A., Morrison, J. G., & Hitchcock, E. M. (1993). Effects of allocation and partitioning strategies of adaptive automation on task performance and perceived workload in aviation relevant tasks. In R. S. Jensen & D. Neumeister (eds.), *Proceedings of the 7th*

International Symposium on Aviation Psychology (pp. 150-155). Columbus, OH: Department of Aviation, The Ohio State University.

Hancock, P. A., & Chignell, M. H. (1988). Mental workload dynamics in adaptive interface design. *IEEE Transactions on Systems, Man and Cybernetics*, *18*, 647-658.

Harris, W. C., Goernert, P. N., Hancock, P. A., & Arthur, E. (1994). The comparative effectiveness of adaptive automation and operator initiated automation during anticipated and unanticipated taskload increases. In M. Mouloua & R. Parasuraman (eds.), *Human performance in automated systems: Current research and trends* (pp. 28-32). Hillsdale, NJ: Lawrence Erlbaum.

Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (eds.), *Human mental workload* (pp. 139-183). North-Holland Elsevier Science: Amsterdam.

Hilburn, B., Jorna, P., Byrne, E. A., & Parasuraman, R. (1997). The effect of adaptive air traffic control (ATC) decision aiding on controller mental workload. In M. Mouloua & J. M. Koonce (eds.), *Human-automation interaction: Research and practice* (pp. 84-91). Mahwah, NJ: Lawrence Erlbaum Associates.

Hilburn, B., Molloy, R., Wong, D., & Parasuraman, R. (1993). Operator versus computer control of adaptive automation. In R. S. Jensen & D. Neumeister (eds.), *Proceedings of the 7th International Symposium on Aviation Psychology* (pp. 161-166). Columbus, OH: Department of Aviation, The Ohio State University.

Kaber, D. B. (1997). *The Effect of Level of Automation and Adaptive Automation on Performance in Dynamic Control Environments* (Tech. Work. Doc. No. ANRCP-NG-ITWD-97-01). Amarillo, TX: Amarillo National Resource Center for Plutonium.

Kaber, D. B., & Endsley, M. R. (1995). *Multitask*©. Lubbock, TX: Texas Tech University.

Kaber, D. B., Endsley, M. R., & Onal, E. (2000). Design of automation for telerobots and the effect on performance, operator situation awareness and subjective workload. *Human Factors and Ergonomics in Manufacturing*, *10*(4), 409-430.

Kaber, D. B., & Riley, J. (1999). Adaptive automation of a dynamic control task based on secondary task workload measurement. *International Journal of Cognitive Ergonomics*, *3*(3), 169-187.

Kessel, C. J., & Wickens, C. D. (1982). The transfer of failure-detection skills between monitoring and controlling dynamic systems. *Human Factors*, *24*(1), 49-60.

Lorenz, B., DiNocera, F., Rottger, S., & Parasuraman, R. (2001). The effects of level of automation on the out-of-the-loop unfamiliarity in a complex dynamic fault-management task during simulated spaceflight operations, In *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting* (pp. 44-48). Santa Monica, CA.

Moray, N. (1986). Monitoring behavior and supervisory control. In K. R. Boff, L. Kaufmann, & J. P. Thomas (eds.), *Handbook of perception and human performance: Volume II: Cognitive processes and performance* (pp. 40-1 - 40-51). New York: John Wiley & Sons.

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

Neter, J., Wasserman, W., & Kutner, M. H. (1990). *Applied Linear Statistical Models* (3rd ed.). Homewood, IL: Irwin.

Norman, D. A. (1989). *The problem of automation: Inappropriate feedback and interaction not overautomation* (ICS Report 8904). La Jolla, CA: Institute for Cognitive Science, U. C. San Diego.

Ntuen, C. A. & Park, E. H. (1988). Human factors issues in teleoperated systems. In W. Karwowski, H. R. Parsaei & M. R. Wilhelm (eds.), *Ergonomics of hybrid automated systems I* (pp. 230-210). Amsterdam: Elsevier.

Parasuraman, R. (1993). Effects of adaptive function allocation on human performance. In D. J. Garland & J. A. Wise (eds.), *Human factors and advanced aviation technologies* (pp. 147-158). Daytona, Beach, FL: Embry-Riddle Aeronautical University Press.

Parasuraman, R., Bahri, T., Deaton, J., Morrison, J., & Barnes, M. (1992). *Theory and design of adaptive automation in aviation systems* (Progress Rep. No. NAWCADWAR-92033-60). Warmister, PA: Naval Air Warfare Center, Aircraft Division.

Parasuraman, R., Mouloua, M., & Molloy, R. (1996). Effects of adaptive task allocation on monitoring of automated systems. *Human Factors*, *38*(4), 665-679.

Parasuraman, R., Mouloua, M., Molloy, R., & Hilburn, B. (1993). Adaptive function allocation reduces performance costs of static automation. In R. S. Jensen & D. Neumeister (eds.), *Proceedings of the 7th International Symposium on Aviation Psychology* (pp. 178-185). Columbus, OH: Department of Aviation, The Ohio State University.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse and abuse. *Human Factors*, *39*(2), 230-253.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model of types and levels of human interaction with automation. *IEEE Transactions on Systems, Man and Cybernetics*, *30*(3), 286-297.

Pope, A. T., Comstock, R. J., Bartolome, D. S., Bogart, E. H., & Burdette, D. W. (1994). Biocybernetic system validates index of operator engagement in automated task. In M. Mouloua, & R. Parasuraman (eds.), *Human performance in automated systems: Current research and trends* (pp. 300-306). Hillsdale, NJ: Lawrence Erlbaum.

Rouse, W. B. (1977). Human-computer interaction in multi-task situations. *IEEE Transactions on Systems, Man and Cybernetics*, *SMC*-7, 384-392.

Rouse, W. B. (1988). Adaptive aiding for human/computer control. *Human Factors*, *30*, 431-438.

Rouse, W. B., Geddes, N. D., & Curry, R. E. (1986). An architecture for intelligent interfaces: Outline of an approach to supporting operators of complex systems. *Human Computer Interaction*, *3*, 87-122.

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors*, *37*(1), 5-19.

Scallen, S., Hancock, P. A., & Duley, J. A. (1995). Pilot performance and preference for short cycles of automation in adaptive function allocation. *Applied Ergonomics*, *26*, 397-403.

Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (eds.), *Human performance in automated systems: Theory and applications* (pp. 37-64). Mahwah, NJ: Lawrence Erlbaum Associates.

Sheridan, T. B. (1997). Supervisory control. In G. Salvendy (ed.), *Handbook of human factors* (pp. 1295-1327). New York: John Wiley & Sons.

Sheridan, T. B., & Verplank, W. L. (1978). Human and computer control of undersea teleoperators. Cambridge, MA: MIT Man-Machine Laboratory.

Wickens, C. D. (1992). *Engineering Psychology and Human Performance* (2nd ed.). New York: HarperCollins.

Wickens, C. D., & Kessel, C. (1979). The effects of participatory mode and task workload on the detection of dynamic system failures. *IEEE Transactions on Systems, Man and Cybernetics*, *SMC*-9, 24-34.

Wickens, C. D., & Kessel, C. (1981). Failure detection in dynamic systems. In J. Rasmussen & W. Rouse (eds.), *Human detection and diagnosis of failures* (pp. 282-293). New York: Plenum.

Wiener, E. L. (1988). Cockpit Automation. In E. L. Wiener, & D. C. Nagel (eds.), *Human factors in aviation* (pp. 433-459). San Diego, CA: Academic Press.

Wiener, E. L. & Curry, R. E. (1980). Flight-deck automation: Promises and problems. *Ergonomics*, *23*, 95-1011.

Young, L. R. A. (1969). On adaptive manual control. *Ergonomics*, *12*(4), 635-657