INDIVIDUAL PREFERENCES USING AUTOMATION

by

JENNIFER E. THROPP

B.S., University of Central Florida, 2001
M.S., University of Central Florida, 2004

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Psychology
in the College of Arts and Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2006

Major Professor: Peter A. Hancock, Ph.D.

ABSTRACT

As system automation increases and evolves, the intervention of the supervising operator becomes ever less frequent but ever more crucial. The adaptive automation approach is one in which control of tasks dynamically shifts between humans and machines, being an alternative to traditional static allocation in which task control is assigned during system design and subsequently remains unchanged during operations. It is proposed that adaptive allocation should adjust to the individual operators' characteristics in order to improve performance, avoid errors, and enhance safety. The roles of three individual difference variables relevant to adaptive automation are described: attentional control, desirability of control, and trait anxiety. It was hypothesized that these traits contribute to the level of performance for target detection tasks for different levels of difficulty as well as preferences for different levels of automation. The operators' level of AC was inversely proportional to LOA preferences. The effects of sensory modality were also assessed. It was also found that operators of both low and high AC levels can adapt to both low and high LOAs in terms of target detection performance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Many jobs which were once completed by means of human skill and cognition have been allocated to machines for a number of reasons; automation serves humans well by doing deeds that are dangerous, excessive, time- and effort-consuming, difficult, or impossible according to human standards. Many attempts at automation have aimed toward taking the human out of the control loop or relegating the human to a supervisory role in the automation loop (Endsley, 1995a). Yet, even with artificial intelligence, machines are indeed fallible and complimented by human operators.

Humans are unlikely to become obsolete in lieu of many completely automated systems, such as the highly unlikely prospect of a fully-automated commercial aircraft (Edwards, 1988). The advantages of the human operator, such as complex pattern recognition and recall from past events to solve new problems, are difficult for a machine to replicate. Even if artificial intelligence is implemented in attempt to replace the human operator, it will require human supervision and management to ensure continued system safety. The human will continue to be an integral part of automation, and so understanding and predicting his or her behavior, abilities, and limitations in these environments will enhance safety, performance, efficiency, and the work experience. Furthermore, these specific needs and preferences among operators can manifest themselves in cognition and personality.

It is often challenging to properly allocate a task to human or machine responsibility, notably if there is a lack of understanding regarding the human's abilities in a task-specific domains and workload levels (Hancock & Scallen, 1996). To further clarify an accommodating allocation, awareness and consideration of the operator's skills, attentional behavior, and resilience to high mental demand could be beneficial. For instance, situation awareness (SA) errors have been attributed to unavailability of appropriate data because of failures in design, human workload, user memory, and operator attention (Ensley, 1995). Thus, safety can be improved in many human-machine environments, such as commercial aviation which depends upon pilot performance (Endsely, 1995).

Individual differences play a dominant role in determining how well a job can be performed, accounting for more variability in performance than differences in systems designs and/or training procedures (Egan, 1988). Yet, many computerized systems have been developed without individual differences in capability. Current technologies in adaptive and flexible automation allow for certain types of individuation based on customization (see Williges, Williges, & Fainter, 1988; Tso et al., 2003). Thus, the potential exists to extend this to a broad range of automated systems, with the inclusion of adaptation based upon cognitive ability as well as preference.

Recent literature reviews regarding the influence of individual differences in operators of automated systems reveal that there is a paucity of theory and empiricism in this domain (Thropp, Oron-Gilad, Szalma, & Hancock, 2004). While a great deal of literature has been published regarding automation and individual difference variables separately, studies that integrate these two concepts are scarce. The reasoning for this may be as simple as a lack of awareness of the potential ways in which human characteristics impact their performance in automated systems.

However, the potential for integrating these two topics is burgeoning. Adaptive automation that functions according to user individuation has many potential benefits, and is an area that warrants investigation. According to Hammer and Small (1995), “an examination of how humans decide to share tasks and information may be a more fruitful area in which to develop a theory of human-associate interaction”. This reflects the sentiment that the field of human factors could benefit by expanding its research into the characteristics of human operators of adaptive automation systems. A strong understanding of human individual differences can therefore serve as an important foundation for developing an adaptive, user-centered form of automation that best accommodates human-machine interaction. This series of experiments comprise one such venture into understanding the role of operator traits in the sharing of information and tasks, as well as the future of adaptive automation based upon individuation.

CHAPTER 2: LITERATURE REVIEW

Development and Nature of Adaptive Automation

Automation actively selects data, transforms information, makes decisions based upon that information, enacts controls processes, and seeks to sustain and enhance system performance while retaining safe operation (Lee & See, 2004). It has recently been defined as the mechanical or electrical accomplishment of work (Wickens & Hollands, 2000), implying that human effort can be replaced by machine action.

‘Work’ in this sense is thus broadly characterized by tasks that can be completed by human, by machine, or by means of collaboration between both parties. ‘Allocation’ is based on various characteristics of each party’s respective effort, such as the environmental needs of the mission and process demands, with the ultimate goal of optimizing task performance.

The first generations of automation were fixed in a static form of task allocation, in which there was a one-time designation for either human or machine task completion (Hancock & Chignell, 1989; Rouse, 1988). Conversely, flexible allocation, in which task responsibility can be switched between parties, suggested that adaptability would be a more effective approach. In adaptive systems, task allocation is dynamic and reflects changing demands posed by the task on the system and the operator (Hancock & Chignell, 1989). For instance, a decline in operator performance can induce the automation to take control of the task in order to avert significant performance failure such as an accident. This type of scenario is made possible by an adaptive system that responds to the human; a static system, on the other hand, would not intervene, and the unfortunate performance failure may result. Thus, the system adapts to the needs of the user, rather than the user adapting to the ways of the system (Edmonds, 1981; Mason & Thomas, 1984). Furthermore, the responsive nature of adaptive automation places greater concern on preventing task overload or underload on the human, making it representative of a human-centered philosophy (Hancock & Scallen, 1996).

Uses of Automation

Automation serves three crucial purposes by performing tasks that are either i) beyond human ability, ii) executed poorly by humans, and/or iii) considered excessively demanding for human operators (Wickens 1992; Lee and See, 204). Thus, these tasks can be generally categorized as unfeasible according to human standards, or feasible yet undesirably taxing.

Automation can be used to complete tasks that human operators cannot do because of limitations on human skills such as complex, time-sensitive mathematical computations and precise coordination of multiple systems. It can also perform activities that require a degree of detail that are beyond human sensory thresholds, such as target detection involving visual stimuli that appear too quickly for detection by the human eye. Also, automation can perform tasks which are unsafe, such as handling hazardous materials or missions that are remotely allocated, as seen in the applications of unmanned aerial vehicles (Mouloua, Gilson, & Hancock, 2003). It can also serve as a peripheral aid for these tasks in which human performance is limited. For instance, it can serve as a memory aid for the human operator who has limited working memory capacity

Human operators may allocate a job to a machine if it is considered undesirable for any reason. For example, if a high level of workload might be imposed by multitasking requirements, an automated system may be used to take control of the excess tasks, perhaps with the use of artificial intelligence or expert systems (Madni, 1988).

Financial interests are yet another possible reason for automation implementation. Machine-operated systems can often times be more cost-effective than their human counterparts. This can be seen in situations where the costs of training and paying human employees to do monotonous work may be excessive. Replacing human workers with automated ones does, however, come with the potential lack of *user friendliness*, in which the machine is unable to replicate the social adaptability of the human operator (Landauer, 1995).

Levels of Automation

Depending on the particular system of interest, automation can be characterized along multiple dimensions. While factors such as triggers, timing, quantity, and quality of information displayed are important, the variability of automation levels is the focus of this paper due to its proposed relationships with operator variables.

A task can be executed by the operator in a completely manual fashion, that is, initiated and controlled through solely human efforts and without machine intervention. On the other end of the spectrum is the case of full automation in which the machine initiates and executes the entire task. This binary arrangement of purely automatic versus purely manual control is the most simple allocation design (Scerbo, 1996). Yet, in between these two extremes are various degrees of partial human versus partial machine control, in which the two parties collaborate on completion of the task. More complex systems may provide these multiple levels of automation (LOAs), varying along a continuum ranging from fully manual to fully automatic control. Furthermore, in flexible and adaptive systems, the levels can change dynamically as functions of environmental demand and system state.

There are many models of automation levels, yielding a varying number of LOAs, (see Verplanck, 1978; Endsley, 1987; Ntuen & Park, 1988). In one such model, Sheridan and Verplank (1978) described a 10-point scale of LOAs, in which higher levels are characterized by an increasing degree of machine control and a consequently decreasing degree of human control (see Figure 1). More specifically, in lower LOAs, the computer makes suggestions to the operator (along with alternative suggestions in the lowest levels), who then selects one of the options. In higher LOAs however, the trend gradually becomes one in which the computer makes decisions with fewer opportunities for the operator to intervene.

High
10. The computer decides everything, acts autonomously, ignoring the human,
9. informs the human only if it, the computer, decides to
8. informs the human only if asked, or
7. executes automatically, then necessarily informs the human, and
6. allows the human a restricted time to veto before automatic execution, or
5. executes that suggestion if the human approves, or
4, suggests one alternative
3. narrows the selection down to a few, or
2. The computer offers a complete set of decision/action alternatives, or
Low
1. The computer offers no assistance: human must make all decisions and actions.

Figure 1. 10-point scale of LOA (Sheridan & Verplank, 1978).

The input functions which make output possible, however, may also be influenced by automation, as described in the four-stage model of information processing proposed by Parasuraman et al. (2000). Sensory processing is the first stage of this model, in which information is perceived and pre-processed. Next, information is consciously processed in working memory. Decision-making can then occur in the third stage, which then dictates the action or behavior in the fourth stage. These four stages, however, often overlap with each other depending upon the nature of the task (Wickens & Hollands, 2000).

Likewise, automation of activity can be described in terms of a comparable four categories of function: information acquisition, information analysis, decision and action selection, and action implementation. Each of these four categories can proceed at different LOAs at any given point in the task (e.g., low LOA on information acquisition, with high LOAs on the other three categories). In the first stage, information acquisition automation, input data is detected and registered by the system. Information filtering may also occur in order to retain and display the most relevant data in the environment; here, the quality and quantity of the information display can impact the human's acquisition process. In the next stage, analysis automation, information manipulation and inferential processing occur. The automation may then predict events using algorithms, or context-dependent data presentation. The third stage is decision automation, in which human decision-making can be assisted or replaced, depending upon the LOA. Once a decision is made, action automation can take place, either aiding or replacing human action. The different LOAs at work in each of these stages could be adaptive in nature, accommodating the specific user needs and preferences present at the time of operation.

There may also be LOA effects upon task difficulty. Dual- and multi-task performance involves the demands of completing two or more tasks at the same time, and is often a performance requirement in many complex and information-rich systems. In the aviation domain, for instance, the pilot may be taxed with simultaneously maintaining a specific altitude while monitoring a radar display. As attention is a limited resource (Navon & Gopher, 1979), performance declines as resources are allocated to multiple tasks at the same time. Should one of the concurrent tasks become more automated, however, performance on the manually controlled task may rebound as resources have been freed to serve the manual task (Hancock, manuscript submitted). Various LOAs can therefore result in different levels of task difficulty, allowing for improved multitask performance as automation assumes more control from the operator.

Changes in the LOA of the task underway can be user- or system-controlled. Automation implemented by the human may be effective, as performance-based data has indicated (see Hilburn et al., 1993). For instance, Harris, Hancock, Arthur, & Caird, (1991) demonstrated that operators more effectively executed a resource management task when they had control over the invocation. However, it may be more beneficial to let the system invoke automation when the operator must engage the automation at a precise time or is too preoccupied to do so (Sarter & Woods, 1994b; Wiener, 1989). The operator's ability to assess the need for automation invocation may also be questionable (Morrison & Rouse, 1986), especially in times of fatigue, when automation may be engaged inadequately (Harris, Hancock, & Arthur, 1993). Finally, in high-risk scenarios, there may be a justification for the automation to assume authority and change the operations underway, despite a lack of confirmation or invocation from the user (Scerbo, 1996).

Thus, it is important to consider the LOAs which may be beneficial for the user. Among factors that may influence the appropriate LOAs are the unique preferences and abilities for each operator. An understanding of these individual differences can aid greatly in designing the most effective interface for each unique operator in various environments and levels of cognitive demand.

Advantages of Automation

Automation provides benefits such as reducing the number of tasks humans must do and attenuating the variability of human performance (Scerbo, 1996). In a dual-task study conducted by Prinzel et al. (2003), physiological evaluations from P300 amplitude measures showed that adaptive automation can improve allocation of attentional resources to the primary task, freeing additional resources for deployment to the secondary task. Thus, operators using adaptive allocation reported the lowest level of subjective mental workload. These results demonstrate that performance can be improved and workload can be reduced when manual and automatic task mode allocations are utilized appropriately.

The use of automation can also reduce sources of human error. As Scerbo (1996) pointed out, human performance is, in part, characterized by variability and inconsistencies which could serve as sources of error. Automated task performance, however, is more likely to be consistent. Allocating jobs to computers and machines also permits multitasking, as the human can embark on additional tasks once others are automated. More tasks can be undertaken when appropriate, as automation can increase the number of systems that can be controlled simultaneously (Sarter & Woods, 1994).

The specific level of control can also be flexible and efficient in adaptive automation. It permits the operator to exercise more or less machine control by changing LOAs as needed (Scerbo, 1996). Finally, automation can be implemented in a variety of settings, from simple household appliances that operate on timers, to more complex systems such as aircraft that maintain and change altitude. According to Parasuraman, Sheridan, and Wickens (2000), there are few systems that cannot be automated to some degree. Thus, there are domain-specific benefits as well. For instance, within the most evidently exploited domain of aviation, pilots experienced reduced flight times, increased fuel efficiency, more effective navigation, and improved perceptual and cognitive activities as a result of automation invocation (Wiener, 1988).

Disadvantages of Automation

While automated systems have demonstrated numerous advantages, there are also problems which must still be addressed. First, there are disadvantages associated with removing humans from the ‘loop’ of the operational environment and the consequential situation awareness (SA) decrements (Sarter & Woods, 1994). To compensate, supervisors may require methods of coping with the volume of environmental events in increased system complexity. Here, a higher LOA would be a likely coping mechanism due to its capacity to use automation while preserving human-in-the-loop benefits.

There may be a degradation of the human skill which is not practiced because the particular task to which it applies has become automated. For instance, manual skills are subject to atrophy without regular use (Wickens, 1992). An automation deficit may occur after a period of automated control, whereby manual performance degrades (Morrison & Rouse, 1986). With an automated system, humans are often delegated to a monitoring task, yet it has been documented that humans are not well suited for extended periods of information monitoring (Parasuraman, 1986; Warm, 1984). Indeed, with a purely automated task, monitoring behavior eventually suffers (Parasuraman, Molloy, & Singh, 1993).

Although automating a task has often been associated with a reduction in human workload, the contrary can also occur, and it may in fact become a burden under periods of high workload (Sarter & Woods, 1994). Part of this problem stems from new demands upon the operator to coordinate and supervise the automation (Woods, 1996). The relationship between workload and task criticality was examined by Olson and Sarter (2000); their results indicated that as system control increased, so do the cognitive demands and the need to interact with the automation.

Confidence in the automated system may also influence the operator's control choices (Lee & Moray, 1992; Muir, 1987) and the assessment of system reliability (Rouse, 1991). While confidence is necessary for the user's acceptance of automation, inflated confidence brought about by excessive trust in automation may induce over-reliance, and finally complacency (Parasuraman et al., 1993). In this situation, operators begin to rely on the automation to the extent that they become less apt to monitor and evaluate the system's performance. Inappropriate conditions of automation use can also occur. When interacting with automation, some individuals practice misuse by violating critical assumptions about the functions and abilities of the system, and also by relying on automation in inopportune situations (Parasuraman & Riley, 1997). Conversely, they engage in disuse when they fail to accept its ability and utility.

Imperfect machine reliability may undermine operator trust, however. The operator may also influence automation, as seen in the role of confidence (Lee & Moray, 1992; Muir, 1987) and experience with the system to assess its reliability (Rouse, 1991). According to Lee (2004), trust can be afforded by the three bases of performance (reliability, predictability, and ability), process (appropriateness of the automation's behaviors), and purpose (whether the automation is functioning as it was intended). Some operators may disuse an automated target detecting mechanism which is known to occasionally err, for instance.

A final disadvantage is that fixed automation can induce LOA changes without consideration of other important variables, such as the environment in which it is operating, task-specific demands, or the specific needs of each individual operator. This may occur simply because the system is unaware of this information. As a result, various needs of the operator are not met, thereby impairing user performance, safety, and positive work experience due to high mental workload and frustration. Adaptive automation can aid in minimizing this rigidity by changing its characteristics in response to the evolving demands as signaled by various triggers. However, there are also many needs which are unique to each operator and cannot always be addressed via traditional triggering mechanisms. These individual differences in operators can influence human performance, and when accommodated, positively impact the human-machine interaction.

Attention and Performance in the Context of Automation

Task performance can be subject to resource limitations (Matthews et al., 2000), especially when the task is very difficult and it consumes a large portion of attentional capacity. However, if more resources are allocated to the task, performance may consequently improve. At a data-limited point in the resource-performance relationship, investing additional effort no longer yields performance gains. Data limits occur in situations where the information quality is poor, for instance, when signals are low-intensity. Likewise, when resources are added or removed, performance may not change in cases of resource limitations. Automation can be implemented to aid the operator in high workload conditions, or when the task difficulty is beyond the operator's ability (Wickens, 1992; Lee & See, 2004).

Time Sharing and Divided Attention

When two tasks or sources of stimuli are presented concurrently, attention must be allocated between them if the individual wishes to attend to them simultaneously. When resources are efficiently allocated between concurrent tasks, time-sharing is facilitated and dual-task performance is enabled (Gopher, 1991). Efficient time-sharing is necessary for all stages of information processing, ranging from sensory perception to working memory to decision and action. Allocation skill can be further improved by training individuals to control attention dynamically in multitask environments (see Gopher, Weil, & Seigel, 1989; Kramer, Larish, & Strayer, 1995). Thus, time-sharing performance can become more efficient with time and experience.

Divided attention can also be guided by specific allocation schedules based upon the relative prioritization of the two tasks; the more important task should receive more attentional resources. Dual-task interference occurs when two or more concurrent tasks are not performed as well as they would be if each was performed alone. This performance decrement is evidence of imperfect time-sharing between the tasks.

Dual-task performance can be undermined by its inherent cognitive demands. Switching attention between tasks can engender a cost in performance (Moray, 1986; Sheridan, 1972), and the operator may thus be inclined to spend more time at a lower priority task rather than exert the effort to switch to the more important task. Furthermore, 'planning' the strategies for task management and attentional allocation schedules are an additional source of cognitive workload; thus, planning may be neglected and the result is a less-than-optimal allocation (Tulga & Sheridan, 1980). Should the tasks be of moderate workload, however, it is more probable that there will be enough remaining cognitive resources to devote to such strategic allocation planning and task management (Hart & Wickens, 1990). Automation can aid in task management by monitoring performance and informing operators of critical events when operators are faced with high workload (Funk & McCoy, 1996; Wiener & Curry, 1980, Hammer, 1999).

When the operator cannot effectively process two tasks concurrently, switching attention between them becomes necessary. The constraints limiting the ability to process two tasks concurrently can involve the channels of the tasks themselves. Perfect time-sharing between two perceptual tasks is often not possible due to limited perceptual resources (see Wickens & Hollands, 2000). There are also limited resources available for response selection. Multiple perceptual and response demands may result in a competition for these limited resources. Interference between the tasks will therefore occur unless one or both are automated (Wickens & Hollands, 2000).

Attention in the Auditory and Visual Modalities

Environmental stimuli can be grouped into various channels in which events may occur, such as visual, auditory, and haptic channels. By sampling these channels for events, one can then gain an understanding of the situation at hand. Sampling rates can be influenced by factors such as the observer's mental model or expectancies of the environment (Bellenkes, Wickens, & Kramer, 1997) and the observer's tendency to match sampling rate with the event rate. Appropriate scanning is often a function of experience with the task at hand (Kundel & LaFollette, 1972). Additionally, observers usually visually fixate on areas of high information or where important information is expected to occur (Yarbus, 1967). For instance, pilots tend to fixate longer on critical instruments (Harris & Christhilf, 1980). Longer fixations are also made upon stimuli with less intelligible meaning, such as displays that are difficult to interpret or have large amounts of information (Mackworth, 1976).

Eye movements are driven by the need to attend to a stimulus in pursuit of information, a process known as visual sampling. In target search sampling, the visual field is scanned for a potential stimulus of interest in an unknown location. In supervisory control sampling, attention is allocated to a specific region in which the target is known to appear (Liu & Wickens, 1992). For an operator monitoring system instruments for critical events, supervisory control performance is of high importance.

Gauges and displays must be sampled adequately to perceive and process events. Certain display characteristics can also guide visual sampling; bright, colorful, changing, abrupt stimulus onset or blinking stimuli can attract visual attention.

When engineering sounds for an auditory display, it is important that the designer not rely on loudness alone to attract the operator's attention, as startle and annoyance can contribute to stress and jeopardized information processing (Wickens & Hollands, 2000). It may be beneficial to instead capitalize on the spatial dimension of sound to allow the listener to distinguish between different auditory channels. Darwin, Turvey, and Crowder (1972) state that three distinct spatial auditory channels can be processed without distraction simultaneously: one for the left ear, one for the right ear, and one in which a sound is presented with equal intensity to both ears. Although it would be too difficult to conduct semantic processing for all three in parallel, this arrangement would better enable the listener focus on one channel while filtering out the other two. In light of this, avoiding temporal overlap among stimuli in these three channels may allow for processing of all three.

Focused auditory attention can be assessed by different methodologies. In monaural listening, two messages are delivered by headphones with equal relative intensity to both ears (i.e., both ears hear both messages). In dichotic listening, headphones present one message to the left ear and another message to the right ear, and one message is heard in each ear (Cherry, 1953; Pashler, 1998). Dichotic listening generally allows for more effective filtering (Egan, Carterette, & Thwing, 1954). Cherry (1953) found that the ability to shadow speech presented to one ear was unaffected by the presence of unrelated speech in the other ear.

Cross-modality Attention

One way to alleviate the problem of limited resources is by delegating processing demands to multiple sensory modalities; time-sharing can be more efficient when utilizing separate rather than common resources. Cross-modal attention time-sharing is usually more efficient than dividing attention between two channels within the same modality (see Wickens, 1980 for a review). According to Wickens and Hollands (2000), this relationship may also be attributable to the distance between intramodal channels.

There are many events in which attention from two different modalities must occur in parallel. Many preset-day interfaces provide the user with important visual and auditory cues simultaneously (e.g., viewing graphics and hearing sound effects in simulators). Parallel processing is required for dividing attention between these multiple simultaneous sources of input.

Redundant audio and visual coding of a particular stimulus can aid in processing. For instance, clicking on a visual computer desktop icon can open a folder and yield a clicking sound. However, there may be instances in which unassociated visual and auditory stimuli may co-occur and demand operator attention (i.e., separate sources of auditory and visual stimuli). In these cases, the visual channel will often be attended to while the auditory channel will be disregarded. This phenomenon of visual dominance favors visual processing over auditory processing in many time-sharing tasks (Massaro & Warner, 1971). Auditory performance will often be worse than visual performance in these concurrent task requirements.

Auditory signal detection, however, has shown superiority over visual target detection in terms of speed, accuracy, and resistance to a vigilance decrement (Davies & Parasuraman, 1982; Warm & Jerison, 1984; Szalma et al., 2004). The reason may be the coupling differences between the modalities (Hatfield & Loeb, 1968). The source of auditory stimulation is usually in close contact with the ears, as sound can be delivered by headphones or surrounds the ears in the sound field. Auditory processing is thus omnidirectional by nature; sound can be received from any direction, whereas visual processing involves selective attention to scan for information (Wickens & Hollands, 2000). Thus, auditory perception and auditory stimuli are closely coupled (Hatfield & Loeb, 1968). There is looser coupling between visual perception and visual stimuli, however, as the observer can avert his or her eyes and head away from the stimulus, potentially impairing performance. This can be somewhat overcome by requiring the observer to maintain a fixed posture facing the display; however, this requirement may increase workload and effort. Previous literature, however, has not shown a modality effect for perceived mental workload in vigilance tasks (Warm et al., 1996; Szalma et al., 2004). Auditory stimuli can thus be more intrusive and more difficult to neglect. The result is commonly a preference for auditory warnings over visual warnings (Simpson & Williams, 1980; Sorkin, 1987).

Sarter (2000) describes how multisensory displays can aid automated system operators in maintaining awareness of the events that are taking place. A disproportionately large number of these cues are presented visually, such as warning lights or instruments, however, other modalities can be employed, thereby offloading demand upon the visual channel. For instance, an automated agent can provide auditory cues to inform the operator when it has detected a critical event or executed an action. Billings (1997) promotes the idea of human-centered automation design which incorporates a greater range of human senses for information deployment and effective communication within the human-machine team. Sarter suggests that attention can be guided and allocated differentially when various modalities are employed. The omnidirectional quality of auditory attention can be advantageous over situations in which visual attention requires location-specific focusing. A three-dimensional audio display space can also direct the operator's attention to a specific location within the system environment, utilizing the spatial aspect of sound. Additionally, sources of auditory cues do not require physical space on a display as would a visually-based instrument.

Working Memory Capacity and Attentional Control

Working memory is used when the task at hand requires both processing and storage of information. However, it has a limited capacity to do so (Tuholski et al., 2001). Consequently, performance may fail when the task demands exceed the working memory capacity (Schweizer, 1996). There may also be individual differences in working memory efficiency which are determined by numerous factors: the processing speed of the central executive, differences in the time information is maintained in the verbal and visual stores, as well as differences in the capacities of the individual stores (Schweizer, 1996).

The capacity limitation is thought to be partially attributable to the extent to which one uses controlled attentional processing (Engle, Kane, & Tuholski, 1999).

Individuals who differ with regard to working memory capacity should then differ when attention must be controlled, a concept which does not apply to automatic processing (Tuholski et al., 2001). According to Engle (2002), working memory capacity may be related to performance on higher-order tasks, as the key is the individual difference in the ability to control attention in order to maintain information in an active and quickly-retrievable state. Thus, attention serves to inhibit or suppress information as needed, and greater working memory capacity refers to a more efficient ability to control one's attention in order to avoid distraction. Accordingly, the individual differences in working memory capacities should be influential in tasks requiring controlled attention.

The verbal component of working memory involves a phonological store for representing information as words and sounds as well as a visuospatial sketchpad for information such as visual images (Baddeley, 1995). Information in the phonological store can be rehearsed vocally and/or subvocally using an articulatory loop. The phonological store therefore involves storing information as a verbal-phonetic code, and the visuospatial sketchpad stores information using a spatial-visual code. The central executive component controls working memory processes and delegates attentional stimuli to either the phonological store or sketchpad. When concurrent activities are underway, these three components may be impaired by interference. With this in mind, one can design a set of tasks in which the processing demands are divided between spatial and verbal resources, rather than tasks that make exclusive spatial or exclusive verbal demands. With two separate stores, the visual and verbal codes do not compete for the same limited processing resources, thereby better enabling time-sharing in a dual-task condition demanding both verbal-audio and spatial requirements. This is supported by previous literature which has often found the strongest effects of working memory to be found between spatial and non-spatial tasks (Daneman & Tardif, 1987; Shah & Miyake, 1996; Smith & Jonides, 1997). Demands upon the central executive can also impair performance, such as strong requirements of controlled attention in activities such as visuospatial rehearsal (Awh, Jonides, & Reuter-Lorenz, 1998). Refining expertise in the particular task domain can reduce memory load, which in turn relaxes demands made upon the central executive, enabling it to adopt additional tasks. This has been demonstrated in experience gained in air-traffic control tasks (Gronlund et al., 1998).

Effective working memory is critical for maintaining awareness of a changing environment. Extra working memory demands from competing concurrent tasks can therefore interfere with situation awareness. Another way to reduce working memory load is by automating a task which would otherwise be cognitively taxing. However, inappropriate use of automation can also lead to a decrease in situation awareness by removing the operator from the loop (Sarter & Woods, 1995).

Stress, Workload, and Automation

Stress is multidimensional in nature (Matthews et al., 1999). The effects of stress on performance can be further specified by considering the different components of information processing which may be impacted (see Hockey, 1986). For instance, stress may affect selective attention and working memory differently. Furthermore, the task itself can be the proximal source of stress, with respect to its various information processing demands (Hancock & Warm, 1989). Therefore, the impact of stress may be task-dependant (Matthews et al., 2000).

Stress can impair working memory such that it can divert attention away from the rehearsal of spatial and phonetic information; anxiety is one such contributor to this degradation (Berkun, 1964). It can also impair situation awareness (Oransu, 1997).

Divided attention studies are often associated with assessments of attentional capacity overload (Matthews, Davies, Westerman, & Stammers, 2000). Mental workload refers to the attentional demands caused by a cognitive task (O'Donnell & Eggemeier, 1986), as well as the degree to which the task demands are effortful and fatiguing (Mulder, 1986).

One of the most traditional and fundamental goals of automation is to reduce the cost associated with human workload, as higher LOAs can generally reduce operator workload (Endsley, 1999). High workload demands and task difficulty can influence the operator to invoke automation (Olson & Sarter, 2000). There are conditions, however, when increasing LOAs do not reduce workload and improve performance in a simple linear fashion.

Automation originally aimed at reducing workload, however, has often resulted in the human being removed the from the control loop and reassigned to the role of system supervisor. At this point, the human is responsible for monitoring the automation (Endsley, 1995a). Unfortunately, humans are often prone to monitoring performance impairments (Davies & Parasuraman, 1982; Parasuramna, 1987; Wiener, 1987).

Extensive manual control can lead to increased errors and fatigue as well as decrease effectives, especially in regards to vigilance and sustained attention (Parasuraman, 1987; Howe, Warm, & Dember, 1995). Sustained attention requires maintenance of focused attention over a relatively long period of time. Vigilance tasks require sustained attention to monitor for specific infrequent events (Davies & Parasuraman, 1982). While many tasks of sustained attention or vigilance are 30 minutes in length or more (e.g., Davies & Parasuraman, 1982), tasks 10 minutes or less can yield similar decrement patterns as a function of time on task (e.g., Craig, Davies, & Matthews, 1987). Vigilance impairments are also exacerbated by increases in task demand, such as multitasking (Parasuraman, Molloy, & Singh, 1993).

Over-reliance on automation can further impair the operator's ability to detect system failures, as information sources are more likely to be ignored (Molloy & Parasuraman, 1997). Vigilance failure is a common component of overreliance (Moiser, Skitka, & Korte, 1994). The vigilance decrement may be reduced if the operator is responsible for monitoring multiple displays in a complex task rather than one display; thus, task complexity may be an important factor (Parasuraman et al., 1993).

Automation can also increase workload if the operator is excessively taxed by the requirement to interact with and supervise it (Olson & Sarter, 2000). Impairments can occur if the cognitive overhead associated with monitoring the system outweighs the benefits of reduced workload (Parasuraman & Riley, 1997). Automation can also increase workload if it is difficult to initiate and engage (Kirlik, 1993; Parasuraman, 2000).

Olson and Sarter (2000) found that aircraft pilots prefer minimal interaction under conditions of high task demand. This may be due to the phenomenon that increased supervision over a system (i.e., having to approve of the automations' suggestions and decisions) can yield an increase in cognitive demands, (Olson & Sarter, 2000). In another study, it was found that automation that adaptively provided decision aids to air traffic controllers only when workload was high was more beneficial than the constant presence of decision aids under both low and high workload levels (Hilburn, Jorna, Byrne, & Parasuraman, 1997). Thus, there is a trade-off between minimizing task-specific attentional demands as well as automation-supervisory demands.

The Role of Individual Differences in Human-Machine Interaction

It is vital to understand the interactive nature of human performance in automated systems, that is, how operators respond to different configurations of the system. Human operators may be more prone to succeed or fail in certain settings, yet this performance variability could result not only from the system, but also from characteristics of the individuals themselves. Performance-based studies have revealed these individual differences among participants with regard to speed and accuracy as well as less obvious factors such as motivation, emotion, and preferences (Matthews, Davies, Westerman, & Stammers, 2000). Understanding the unique preferences and abilities of operators can aid greatly in designing the most effective interface for each operator.

Importance of System Individuation

An understanding of individual differences can be used to anticipate sources of human error, behavior, and personal preferences. Egan (1988) has argued that oftentimes user differences account for much more of the variability in performance than do system design variables. Furthermore, user differences in human-computer settings may account for ten times the amount of variance that is attributed to different training procedures.

Hancock and Scallen (1996) noted that it is inadequate to make a clear-cut decision as to whether a task is more suited for either human or machine allocation. This is because there will typically not be enough knowledge about the situation to make this simple binary division of labor. More specifically, an effective allocation could be based upon the operator's skills, ability to perform under high mental demand, and the nature of the task(s). Thus, an understanding of the operator's individual trait characteristics can add to the knowledge regarding the situation. Here, adaptive task allocation can more accurately gauge the possible task demands to be made upon the operator based upon his or her abilities. According to Scallen and Hancock (2001), traditional static function allocation is faulty in part because the Fitts report (Fitts, 1951), which lists static human versus machine capabilities, failed to consider the context of the application. This context, which refers to the specifics of the task and environment, can also be extended to include individual differences among operators.

To illustrate, Young and McNeese (1995) showed that problem-solving was dependent upon interactions between neurological processes as well as the environmental information perceived by the human operator. In another example, situation awareness (SA) errors have been attributed to unavailability of appropriate data because of failures in design, human workload, user memory, and operator attention (Ensley, 1995).

Adaptive systems may also be able to accommodate individual differences in SA. It has been estimated that poor SA has been at the root of 88% of the major airlines accidents caused by human error (Endsley, 1995). Furthermore, SA errors have been attributed to unavailability of appropriate data because of failures in design, human workload, user memory, and operator attention. Thus, it seems clear that automation which could compensate for limitations of human cognition would significantly aid in SA quality, thereby preventing a large source of accidents and other performance failures.

Many computerized systems have evolved without any consideration of the user's capability, let alone the range of capability of a population of users. However, the consideration of individual differences is worthy for four major reasons:

1) Individual differences play a dominant role in determining how well a job can be performed, accounting for more variability in performance than differences in systems designs and/or training procedures (Egan, 1988). Individuation would allow for increased productivity and more opportunities for different types of people to become users of the system.
2) Personnel selection for certain jobs has often been driven by personality questionnaires (Highhouse, 2002), but this technique would not necessary apply to situations where computerized systems are meant to be used by the public at large. Even in job-related employee selection, many users are excluded due to their individual differences, when instead, individuation could enable more employees to use the systems.
3) Training expenses can be reduced. Rather than training all users to use the system in the same way, the system can accommodate different user coping behaviors. Alternatively, in instances where a common interface configuration is desirable or necessary, one could design training regimens tailored to specific, relevant user characteristics or learning styles.
4) Advances in current technology allow for the accommodation of individual differences in different users (see Schmorrow, 2002). Through prediction and understanding of performance differences, more systems can be developed which support the characteristics of a wider variety of individuals.

Many of today's flexible and adaptive interfaces allow for accommodating individual differences among operators. In the flexible interface, the operator can actively customize the various characteristics of the interface to suit certain needs and/or preferences. In contrast, the adaptive interface changes automatically to suit the user; it is also capable of evolving to meet these needs as they change over time (Williges, Williges, & Fainter, 1988).

Current technologies in adaptive and flexible automation allow for certain types of individuation and customization. In one example of adaptive technology, the Augmented Cognition program of research (Schmorrow, 2002; St. John, Kobus, & Morrison, 2002; Dorneich, Whitlow, Ververs, & Rogers, 2003) aims to enhance human performance through improving the system's dynamic adaptability to the user's current needs. The goal of the program is a closed-loop human-computer system in which the user's real-time state is assessed through sensors and detectors, and then serves as a basis to which the system will adapt. In particular, the system will adapt to the user's current cognitive state, including cognitive workload, arousal, stress, attention, and both spatial and verbal working memory. Adjustments in environment and workload will then automatically occur in response to the cognitive assessment. The applications of this technology are largely military, and directed towards the warfighter's performance when interacting with computers, such as monitoring and process management by providing the correct quantities and qualities of information at the appropriate times.

In another example of customization, an automated UAV system described by Tso et al. (2003) permits operators to determine their preferred level of control in a mission plan. The operator can also choose the amount and nature of information available in the interface, by opting for an instant return of images or task performance results to decide when to request images and in what order they shall be presented. Customization has also been implemented in F-18 aircraft, in which pilots can opt to place a chosen data display on any of the five CRTs available to them (Williges, WIlliges, & Fainter, 1988). Note, however, that these instances of individuation are passive applications, since the designers permit the users to customize the interface according to their subjective preferences. While such efforts are a positive step, individuation can also be improved by considering how operator traits relate to performance with automated systems. The operator may not always be able to select the appropriate LOA for the given scenario, as with the case of automation overreliance, for instance (Moiser, Skitka, & Korte, 1994).

A theoretical and empirical approach to establishing these relations will improve the application of individuation in the design process and provide design criteria other than subjective user preferences. However, such a systematic approach requires identification of relevant variables and how they might relate to various aspects of automation.

Individual Differences Variables to Consider

While it is important to assess human traits in regards to automation, doing so is often challenging for two reasons, according to Sheridan (1988). Firstly, supervisory machine control requires cognition and mental activities which are not directly observable. Rather, they must be inferred, as there are limitations of directly measuring mental events. Secondly, a free interaction between human and machine makes it difficult to distinguish between the underlying human and machine information processing that yield outputs. This arrangement enables two sources of input, but one final performance output. The result is an integration of mental models internal to the human operator and external computer decision aids.

Many individual difference variables should be investigated (Thropp, Oron-Gilad, Szalma, & Hancock, manuscript submitted); the following variables have been chosen based upon previous findings that strongly extend to the adaptive automation domain. The focus of this paper is upon three of the many potentially important variables that likely influence performance on automated tasks.

Attentional Control

Attention sharing can be a critical component of successful performance; however, Sheridan (1988) noted that the human operator may not be able to shift attention between tasks and information sources as quickly as computers. Sheridan also states that attentional task demands have four main attributes: i) specific resources to be assigned to perform specific tasks, including both human resources such as senses, motor capabilities, and memory, and the computer's resources including sensors, decision-making aids, automated control); ii) the amount of time and effort needed for task completion; iii) the amount of time and effort available for task completion; and iv) the reward for successful task completion and/or cost of unsuccessful task completion.

According to Navon and Gopher (1979), attention is a limited resource which can be divided among stimuli and tasks. One important implication of limited attention is that performance can incur decrements if there are not enough resources to meet the demands of the task. In addition, it is likely that individuals inherently vary in the resources they can allocate and also in the efficiency of such allocation.

Attentional control may be a contributor to human-automation interaction in terms of SA. The ability to divide attention has been positively linked to improved SA (Endsley & Bolstad, 1994). O'Hare (1997) found a positive link between better divided attention performance and SA, while Gugerty and Tirre (1997) found SA to be associated with high working memory capacity, visual processing, temporal processing, and time-sharing.

Attentional control is important for the ability to orient one's attention to the appropriate aspect of an environment (Derryberry & Reed, 2002). Attentional focusing and shifting measures have been combined into a new scale of Attentional Control (Derryberry & Reed, 2002), which measures a general ability to orient one's attention to the appropriate aspect of an environment as necessary. The Attentional Control Scale has specific subfactors for focusing attention, shifting attention between tasks, and controlling thoughts.

Attentional allocation may also correlate with other individual differences. For instance, aspects of attentional control have been described as a function of personality (Derryberry & Reed, 2002). There may also be individual differences in the ability to allocate these attentional resources (Norcio & Stanley, 1989), and adaptive systems can address this by parsing tasks into separate windows within the interface. With each window devoted to a particular task and information type, the user with good attentional allocation can keep track of these different tasks and thereby complete multiple tasks simultaneously. Conversely, the individual with poor allocation skills may have trouble managing parallel windowing and prefer information displays that present data in a singular fashion. Those with high attentional control may more successfully attend to task-relevant stimuli and exhibit higher performance as attentional demands increase.

High attentional control may therefore be conducive to more manual control of system functions, although this highly adaptive characteristic would likely enable the individual to operate successfully at any LOA. Operators with lower attentional control may be facilitated by higher LOAs, as the operator may have deployed these resources to other environmental stimuli instead of task-relevant stimuli. In higher LOAs, the machine may also provide messages indicating its decisions and actions; this may prompt operators with low attentional control to redirect their attention to the appropriate portion of a display.

The Desire for Control

One's desire for control refers to the motivation to control the events in one's life, and relates to the desire to be a leader and make one's own decisions (Burger & Cooper, 1979; Bubb-Lewis & Scerbo, 1999). The desire for control is assessed using the Desirability of Control Scale (Burger & Cooper, 1979), a reliable and valid measure with high internal consistency (McCutcheon, 2000). Increased desire for control has been associated with decreased feelings of powerlessness and passivity whereas low desire for control has been linked to unassertiveness, feelings of powerlessness, and decreased tendency to try to influence others (Burger & Cooper, 1979). Although desire for control has not been linked to changes in performance on an adaptive communication task

(Bubb-Lewis & Scerbo, 1999), it has not been investigated for its interactions with various LOAs.

Previous work has indicated the important role of feelings regarding control. Kantowitz and Casper (1988) suggested that many pilots value perceived control and dislike the ability for automation to override human control. These findings indicate that some operators may have increased preference for control over automated systems. One way to further examine this relationship would be to assess the operator's desirability of control.

Since different LOAs are possible within a given system, one's desire for control may be more prevalent within certain levels. Those with a high desire for control may prefer to maintain control of certain tasks, rather than accept machine control. These individuals may also have more specific preferences over how these tasks are allocated between themselves and the system. They may prefer a flexible system in which they can intervene to disengage the automated control and assume manual control as they deem necessary. High desirability of control may facilitate tendencies towards selecting lower LOAs wherein the automation proposes action but cannot act without explicit operator consent; this allows the operator to interact more frequently with the system. Conversely, those with low desire for control may more readily accept higher LOAs, in which automation acts without explicit operator consent, demanding less of the operator's input. Thus, desirability of control may be associated with not only a preference for a particular LOA, but it may also be important for controlling LOA changes over time.

Adaptive automation can reflect the level of control chosen by different users, while also taking into consideration other important factors such as workload level, time pressure, task difficulty, and task criticality (Parasuraman & Riley, 1997). Thus, fitting the LOA and management style of the automation to the individual's level of desire for control may reduce the workload and stress associated with task performance. Operators who wish to maintain control over the system may experience reduced stress and workload if they are given the choice of LOA, rather than be confined to a highly automated task which is fixed. Those with low desirability of control may therefore experience less stress and workload at higher LOAs when the machine has more control over the task and system functions. It is possible that those with low desirability of control may be more adaptive to accepting various levels of control over events in general than those who want more control, based on a lack of requirement for control in order for preferences to be satisfied.

Trait Anxiety

Trait anxiety is a facet of the more general trait of neuroticism within the Big Five factor model of personality (Costa & McCrae, 1992b). Impairments related to these traits are strongly linked with impairments on tasks requiring short term memory and attentional resources (Eysenck, 1992; Mueller, 1992). One way to assess anxiety is through the State-Trait Anxiety Inventory (STAI: Speilberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). According to Spielberger's (1972) state-trait anxiety theory, trait anxiety is stable and more pervasive in its effect on performance than state anxiety (Matthews & Deary, 1998).

The effects of anxiety on performance are very task-specific (Matthews et al., 2000; Eysenck & Eysenck, 1985). Tasks that are difficult and make strong demands upon short-term and working memory tend to be among the most vulnerable to the effects of anxiety, along with secondary tasks in dual-task paradigms (Eysenck, 1992; MacLeod & Donnellan, 1993; Sorg & Whitney, 1992; Darke, 1988; Eysenck, 1997). For instance, anxiety has been shown to impair the cognitively demanding nature of military aviation (Bartram & Dale, 1982). Anxious individuals can perform well on difficult tasks if they are given success feedback, as well as being given failure feedback on easy tasks (Weiner & Schneider, 1971; Eysenck, 1981). They also tend to be pessimistic about their future task performance (Eysenck & Derakshan, 1997) and more concerned about their successful performance than task failures (Calvo & Eysenck, 1998). Further, the tendency of anxious individuals to focus attention upon themselves may have a negative impact on performance. When an individual is processing worry, fewer resources are thus devoted to the task at hand, thereby potentially impairing performance (Elliman et al., 1996; Matthews et al., 2002). Anxiety is thought to direct selective attention to threat stimuli (Wells & Matthews, 1994). Eysenck (1992) has also suggested that trait anxiety is supported by a broad attentional focus which narrows to accommodate a threat stimulus.

An automated system may adapt to the user with high trait anxiety by providing additional success and failure feedback; this could be in the form of pop-up messages or auditory cues. These users may also benefit from higher LOAs, especially during times of peak task demand and dual-tasking requirements, cases in which highly anxious users are expected to show performance decrements. Higher LOAs would likely decrease the anxious user's stress and mental workload. Individuals with low trait anxiety, on the other hand, may be more adept at allocating resources to the current task; thereby rendering them more adaptive to increasing task demands. Thus, operators with low trait anxiety may be able to execute more difficult tasks and multiple tasks simultaneously and less needful of high LOAs.

CHAPTER 3: THE PURPOSE OF THE CURRENT STUDY

The Individuation Process

Figure 2 provides an illustration of the individuation process and how it might fit into a general model of automation. The assessment and implementation of trait measures may be incorporated into the adaptive automation paradigm originally proposed by Parasuraman, Sheridan, and Wickens (2000). The overlaid individual differences intervention demonstrates how they can influence the choice of specific LOAs by means of the operator's information processing ability. For instance, trait anxiety and attentional control may determine the appropriate LOA by means of impacting the earlier stages of information processing, acquisition, analysis and decision. Desirability of control may dictate LOA though the action phase of information processing.

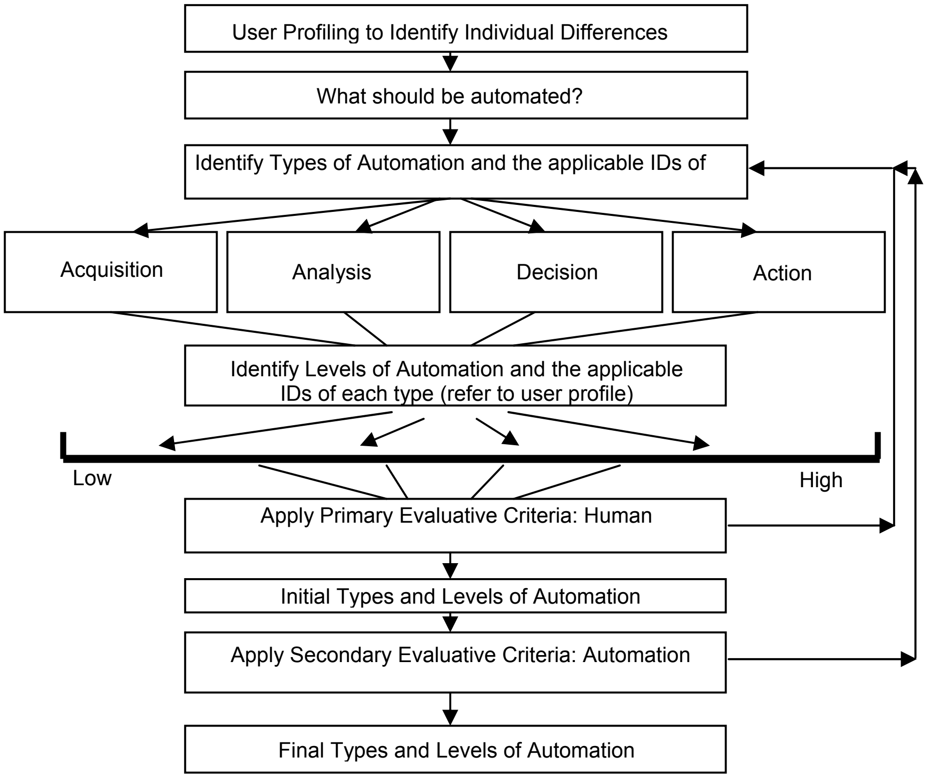


Figure 2. Individual differences assessment overlaid on Parasuraman, Sheridan, and Wickens (2000) model for types and levels of automation.

General Hypotheses

Regarding the role of individual differences and task performance with an automated system, a number of hypotheses emerge.

1. Operators with low attentional control will have worse performance in dual-task conditions due to the increased number of channels to be monitored. This is largely driven by impairments in engaging, focusing, and disengaging from information when appropriate.

    a. However, this dual-task performance decrement may be attenuated if they select higher LOAs to aid them.
    b. Because of the higher task demand in dual-task conditions, they will select higher LOAs. Higher LOAs will be preferred in general.
    c. As a result of selecting higher LOAs for assistance, the dual-task performance decrement may be attenuated.
    d. Workload reports will be higher under dual-task conditions and at lower LOAs. There will be an interaction between LOA and task demand, with higher workload reported for dual-tasking conditions executed at lower LOAs. If higher LOAs are selected, the reported workload will be reduced.
2. Operators with high attentional control have greater ability in engaging, focusing, and disengaging their attention when necessary; this generally renders them more adaptive to various levels of task demand. This will enable them to complete dual-task conditions more successfully compared to those with lower attentional control; however, they will likely also perform single-task conditions just as well if not better than those with low attentional control.
    a. The ability to perform tasks successfully will enable those with high attentional control to chose lower LOAs in both single- and dual-task conditions; however, their adaptability will give them the flexibility to select any LOA. They will prefer lower LOAs in general.
    b. They will report lower overall workload across conditions.

3. Operators with high desirability of control will prefer and select lower LOAs in an automated system. They will want to perform the task as much as possible rather than relegate the task to automated control.
    a. Under conditions of high task demand, as in dual-task conditions, they will continue to prefer lower LOAs. If one task needs to be automated due to high demand, they will still want to approve the machine's actions before they are executed.
    b. They will report lower workload and stress when LOAs are lower, as this will compliment their preferences. When LOAs are higher, increased stress may be driven by increased frustration.
    c. By complimenting their preferences, lower workload may aid performance, and this may improve dual-task performance more than single-task performance due to a possible interaction.
4. Operator with low desirability of control will prefer and select higher LOAs in an automated system; however, they may be more adaptive in that they do not have a control requirement to be satisfied. Thus, they may also choose lower LOAs. They may report less workload under higher LOAs, and increased workload under lower LOAs.
5. Operators with high trait anxiety will select higher LOAs, as they tend to exhibit performance impairments in tasks demanding attentional resources. This is because attentional resources are being deployed to processing worry rather than the current task(s).

    a. This performance decrement will be exacerbated by dual-task conditions, as more attentional demands are imposed.
    b. Dual-task decrements may be attenuated if higher LOAs are employed to aid in performance.
    c. Higher workload will be reported, and will be exacerbated by higher task demand, as in the case of dual-task requirements.
6. Operators with low trait anxiety are more adaptive than those with high trait anxiety, and will therefore exhibit better performance, especially in dual-task conditions.
    a. They will also select from a wider range of LOAs, with a greater tendency towards lower LOAs than those with low trait anxiety.
    b. They will report lower workload in general.
7. Mutitasking conditions will generally yield higher stress and workload reports than single-task conditions; however, these effects may be attenuated if operators select higher LOAs in multitask conditions to alleviate their perceived increase in perceived workload.
8. Operators will be more likely to select higher LOAs in multitask conditions.
9. Auditory task performance may be more accurate than visual task performance due to the closely coupled nature of auditory stimuli and the general superiority of auditory monitoring over visual monitoring.

Table 1 summarizes these hypotheses for each of the three individual differences to be assessed.

Table 1. Summary of experimental hypotheses as functions of traits.

Trait		LOA chosen	Dual-Task performance	Workload
Attentional Control	Low	Higher LOAs chosen	Generally impaired	Lower at high LOAs
	High	Lower LOAs likely, although any level is possible	More successful than those with low attentional control	Generally lower across LOAs
Desirability of Control	Low	Higher LOAs	Possibly improved under high LOAs	Possibly higher at low LOAs
	High	Lower LOAs	Possibly improved under low LOAs	Lower at low LOAs
Trait Anxiety	Low	Lower LOAs	Performance decrement exacerbated by dual-tasking	Generally lower across LOAs
	High	Higher LOAs	More successful than those with low trait anxiety	Lower at low LOAs

These hypotheses were tested in the following series of four experiments. Experiment 1 was conducted to assess the distribution of attentional control, trait anxiety, and desirability of control scores and to recruit participants based upon these traits. Experiment 2 sampled individuals with low and high AC scores and examined their preferences for various LOAs, performance in perceptual tasks, and subjective workload during these tasks. In Experiment 3, performance and subjective workload were again assessed for low and high AC groups; however, they were fixed in low and high LOAs during the perceptual tasks rather than given a choice of LOAs. In Experiment 4, the medium AC group was sampled and performed the same tasks as in Experiment 2, wherein LOA preferences, performance, and workload were measured.

CHAPTER 4: EXPERIMENT 1 – INDIVIDUAL DIFFERENCE SCREENING

Experiment 1 was conducted to assess the distribution of attentional control, trait anxiety, and desirability of control in the population to be sampled. The primary individual difference variable of interest was attentional control, based upon its hypothesized effects upon LOA preferences and task performance with respect to LOA variations. Desirability of control and trait anxiety were also measured within the same sample. The lower and upper quartiles were determined for the distributions of each of these trait variables. Scores on the Automation Level Preference Index (ALPI) were also correlated with each of the trait variables. Correlations between the trait variables were also calculated.

Hypotheses

1. Individuals with low AC will prefer higher LOAs on the Automation Level Preference Index (ALPI), while individuals with high AC will prefer lower LOAs on the ALPI.
2. Individuals with low TA will prefer lower LOAs on the ALPI, while individuals with high TA will prefer higher LOAs on the ALPI.
3. Individuals with low DC will prefer higher LOAs on the ALPI, while individuals with high DC will prefer lower LOAs on the ALPI.
4. AC will be inversely proportional to TA, as reported by Derryberry and Reed (2001).
5. AC will be directly proportional to DC.

6. TA will be inversely proportional to DC.

Methodology

Experimental Participants

Three hundred and three undergraduates (227 females and 76 males) were recruited from undergraduate psychology courses at University of Central Florida in order to fulfill course requirements and received course credit. Their ages ranged from 18 to 39 years, with a mean of 22.25 years (*SD* = 3.52). All participants had normal or corrected-to-normal vision and were free of known hearing impairments.

Materials

All individual differences questionnaires were administered in paper and pencil format. Attentional Control (AC) was assessed using the Attentional Control Scale (Derryberry & Reed, 2002; see Appendix A). Desire for control (DC) was assessed using the Desirability of Control Scale (Burger & Cooper, 1979; see Appendix B). Scores on the DC scale range from 20 (low) to 140 (high). AC scores can range from 20 (low) to 80 (high). Trait Anxiety (TA) was assessed by extracting the relevant items from the State-Trait Anxiety Inventory (Speilberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983; see Appendix C). TA scores can range from 20 (low) to 80 (high). The Automation Level Preference Index (ALPI; see Appendix D) was also administered to give an indication of the degree to which individuals prefer automated versus manual control of a system given a particular scenario of use. ALPI scores are based on a proportion chosen by the participant and can thus range from 0% to 100%.

Experimental Design and Procedure

After a brief description of the experiment by the experimenter, participants were presented with an informed consent form that affirmed they were at least 18 years of age and explained their rights as participants (see Appendix E). All data sheets and electronic data collected were coded with a participant number to maintain confidentiality.

All participants completed the three individual differences measures: the Attentional Control Scale, the Desirability of Control Scale, and the trait anxiety portion of the State-Trait Anxiety Scale. The order of questionnaire administration was counterbalanced across participants.

After completing the trait questionnaires, participants completed the ALPI, in which they will indicate the degree of computer assistance they prefer for a given scenario. They were instructed to read a brief description of a situation involving a target detection task and an automated computer decision aid. They were then asked to choose their preferred proportion of computer assistance they would prefer to invoke to optimize task performance.

Participants then had the option to complete a contact information form in which they gave consent to be contacted for recruitment to subsequent studies (i.e., Experiments 2, 3 or 4; see Appendix F).

Results

A completely within-subjects design was used. Attentional control, trait anxiety, desirability of control, and the Automation Level Preference Index were within-subjects. A summary of the descriptive statistics for each individual difference variable assessed in Experiment 1 can be viewed in Table 2.

Table 2. Descriptive Statistics for Individual Difference Variables.

Variable	Mean	SE of Mean	SD	25th Percentile	75th Percentile
AC	49.58	.528	9.198	44	56
TA	40.53	.498	8.665	34	46
DC	101.35	.805	14.01	93	111
ALPI	65.08	1.045	18.184	50	75

Attentional Control

Measures of central tendency for AC were calculated using SPSS for Windows (version 11). Across all participants, the mean AC score was 49.58 (*SD* = 9.198). The scores ranged from 24 to 74. The distribution was assessed for skewness and kurtosis. With large samples (N>200), visually examining a graphical depiction of the distribution's shape is recommended over using formal tests of skewness and kurtosis (Tabachnik & Fidell, 2001). The distribution of AC scores among all participants appeared symmetrical (see Figure 3). The coefficient alpha was .87.

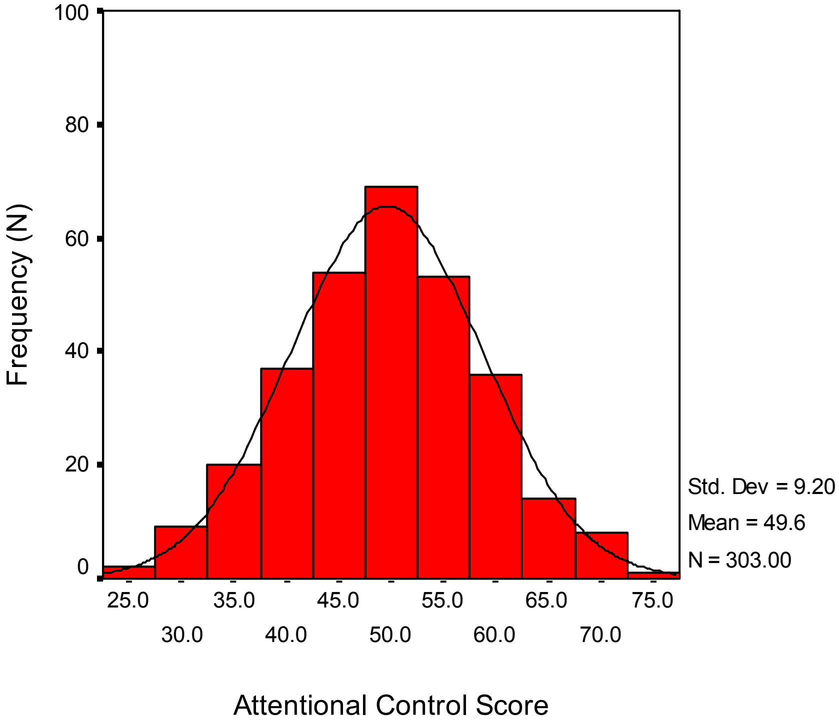


*Figure 3***.** Histogram showing AC distribution among all participants with normal curve overlaid.

To separate 'low' and 'high' AC scores and maximize variance between the groups, a quartile split was performed with SPSS. This procedure indicated that the lower-bound of the upper quartile was an AC score of 56. Conversely, the upper-bound of the lower quartile was an AC score of 44. Thus, the high AC category was determined by a AC scores ranging from 56 to 80, while the low AC scores ranged from 20 to 44 (see Figure 4 for an overlay of the quartile splits upon the distribution). Thus, the range of AC scores within both the low and high AC groups was 24 points.

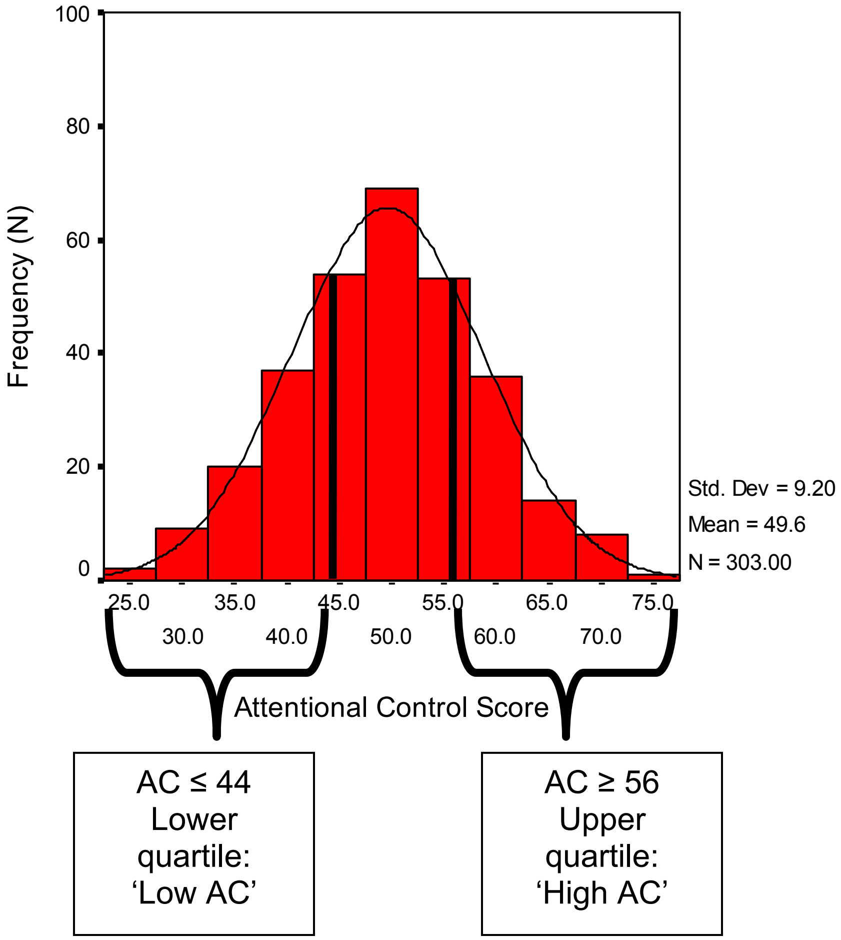


Figure 4. Histogram which features the 'Low AC' and 'High AC' overlay upon the AC distribution.

Trait Anxiety

Measures of central tendency for TA were calculated. Across all participants, the mean TA score was 40.53 (*SD* = 8.665). The scores ranged from 23 to 73. The distribution of TA scores among all participants can be viewed in Figure 5. The coefficient alpha was .87.

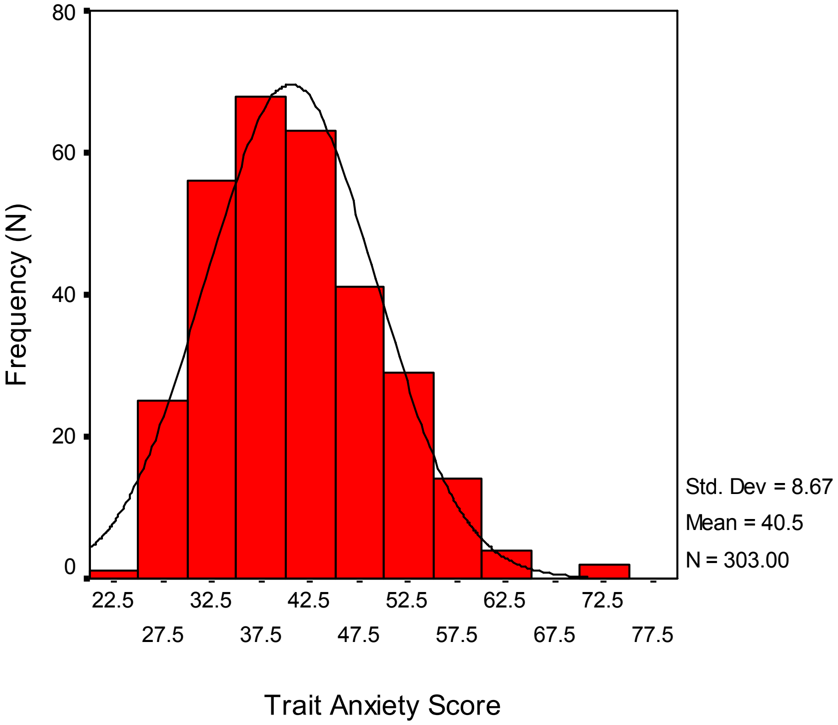


Figure 5. Histogram depicting distribution of TA scores among all participants with normal curve overlaid.

Desirability of Control

Measures of central tendency for DC were calculated. Across all participants, the mean DC score was 101.35 (*SD* = 14.014). The scores ranged from 44 to 134. The distribution of DC scores among all participants can be viewed in Figure 6. The coefficient alpha was .82.

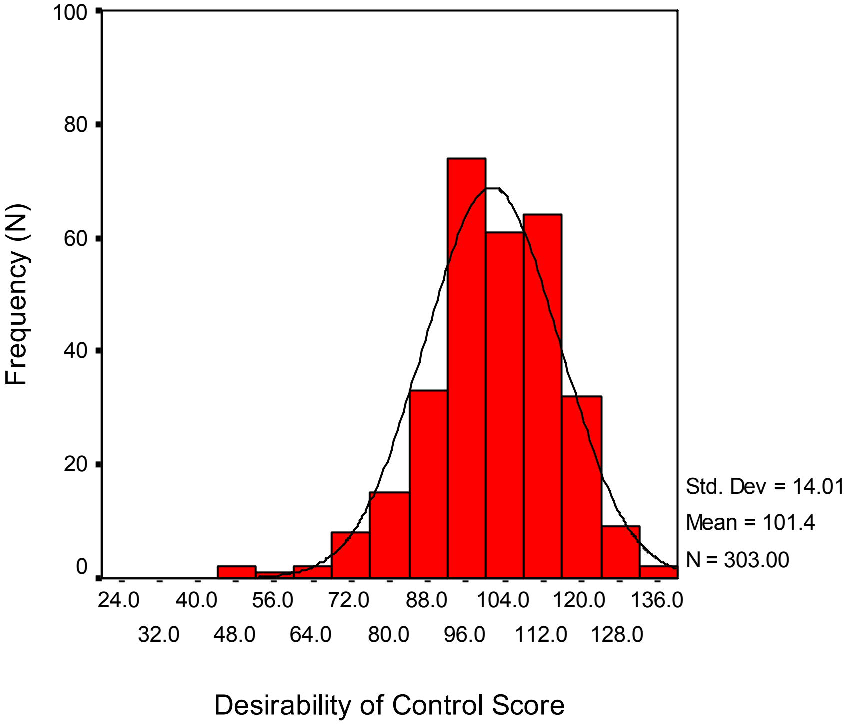


Figure 6. Histogram depicting distribution of DC scores among all participants with normal curve overlaid.

Automation Level Preference Index

Measures of central tendency for the ALPI were calculated. Across all participants, the mean ALPI score was 65.08 (*SD* = 18.184). The scores ranged from 44 to 134. The distribution of ALPI scores among all participants can be viewed in Figure 7.

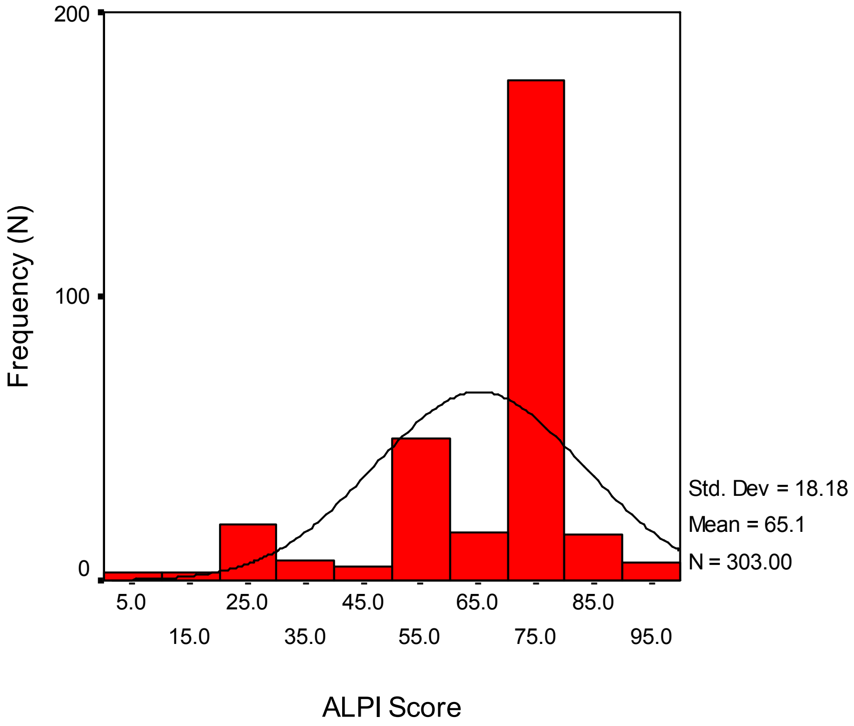


Figure 7. Histogram depicting distribution of ALPI scores among all participants with normal curve overlaid.

The ALPI distributions for the low, medium, and high AC groups were also examined. Among those with low AC, ALPI scores ranged from 11 to 90, with a mean of 67.46 % machine control (*SD* = 16.420); see Appendix AA for the distribution of scores. Among those with medium AC, ALPI scores ranged from 0 to 100, with a mean of 65.58 % machine control (*SD* = 17.606); see Appendix AB for the distribution of scores. Finally, among those with high AC, ALPI scores ranged from 0 to 88, with a mean of 61.53 % machine control (*SD* = 20.606); see Appendix AC for the distribution of scores.

Correlations among individual differences

The Pearson's correlations among TA, AC, DC, and the ALPI can be viewed in Table 3.

Table 3. Pearson's Correlations Among Individual Difference Variables

		TA	AC	DC	ALPI
TA	Pearson correlation	1	-.448*	-.353*	.046
	Sig. (2-tailed)	.	.000	.000	.423
AC	Pearson correlation	.	1	.346*	-.100
	Sig. (2-tailed)		.	.000	.082
DC	Pearson correlation			1	-.099
	Sig. (2-tailed)			.	.084
ALPI	Pearson correlation	.			1
	Sig. (2-tailed)				.

* $p < .01$

Discussion

Experiment 1 was intended to categorize participants into groups of low, medium and high attentional control (AC) scores, as it was hypothesized based on literature by Derryberry and Reed (2002) that AC will have an effect on the ability to focus attention, shift attention between tasks, and control thoughts as necessary. Additionally, attention is a limited resource which can be divided among stimuli and tasks (Navon & Gopher, 1979). Previous investigation has documented that the ability to divide attention has been linked with situation awareness (Endsley & Bolstad, 1994; O'Hare, 1997) as well as temporal processing and time-sharing (Gugerty & Tirre, 1997). As described in the literature review, there are thus many hypothesized relationships regarding the interaction of AC and automation preferences, as well as with task performance. Specifically, it was predicted that those with low AC would prefer higher LOAs, while those with higher LOAs would prefer lower LOAs. Those with low AC would also particularly prefer higher LOAs under increasing task demand, such as dual-task conditions. Furthermore, AC is a relatively uncommonly tested variable and has not yet been experimentally assessed in the domain of human-automation interaction. It was thus chosen as the primary recruitment variable for Experiments 2, 3 and 4.

Desirability of control was also assessed within the same population using the Desirability of Control Scale (DC: Burger & Cooper, 1979). DC refers to the motivation to control events, and the desire to be a leader and make one's own decisions (Burger & Cooper, 1979; Bubb-Lewis & Scerbo, 1999). It was chosen as a potentially relevant individual difference variable due to its potential influence on preferences for an automated agent to assume more or less control of the task at hand. Previous literature has also indicated that feelings regarding control extend to the domain of automation, notably with aircraft pilots (see Kantowitz and Casper, 1988).

Trait anxiety was also measured due to its effects upon short term memory and attentional resources (Eysenck, 1992; Mueller, 1992). The State-Trait Anxiety Inventory (STAI: Speilberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) was administered. It was hypothesized to impact performance with various LOAs due to its influence upon tasks that tax working memory and dual-tasking (see Eysenck, 1992; MacLeod & Donnellan, 1993; Sorg & Whitney, 1992; Darke, 1988; Eysenck, 1997). Specifically, it was hypothesized that participants with high TA would benefit from higher LOAs, especially during times of peak task demand and dual-tasking requirements. Conversely, those with low TA may be more efficient in allocating resources to various concurrent tasks and may prefer lower LOAs.

The coefficient alphas of AC, TA, and DC all indicated high reliability. The distributions of the trait variables scores were also assessed for skewness and kurtosis were visually assessed using a graphical depiction of the distribution's shape (see

Tabachnik & Fidell, 2001). The distribution of AC scores was also the most visibly normal compared to those of the TA and DC scores. The normal distribution of the AC scores suggests that statistical inference would be generally robust (Bradley, 1982) and the logic behind hypothesis testing would be valid as the population from which the sample was drawn was also normally distributed (Field, 2005). The trait anxiety score distribution appeared positively skewed, while the desirability of control scores appeared negatively skewed.

Despite the skewness of DC and TA, these variables did yield significant correlations with AC. The correlation between TA and AC was -.45, whereas Derryberry & Reed (2002) also reported a moderately negative correlation (-.55). This suggests that individuals with low AC tend to have higher TA, and we would anticipate observing results in which those with low AC demonstrate characteristics of individuals with high TA, such as limited attention resources in demanding tasks (Eysenck, 1992; Mueller, 1992), notably dual-task performance (Eysenck, 1992; MacLeod & Donnellan, 1993; Sorg & Whitney, 1992; Darke, 1988; Eysenck, 1997). Those with low AC may also therefore strongly benefit from higher LOAs due to this correlation with TA; this applies to both performance and subjective workload reports. AC also correlated with DC (.35), suggesting that individuals with low AC also have lower DC and will thus be inclined to lower LOAs, and high AC individuals will tend to have higher DC and thus prefer higher LOAs.

None of the individual difference variables significantly correlated with the ALPI. The participants' mode response on the ALPI was to give the computer 75% of the control, wherein the computer made all detection decisions, and the human operator retained 25% of the control by supervising the computer's decisions and reserving the right to veto when the computer erred. This mode response also indicates the participants' approval of the computer's 90% reliability.

Implications for Results

The mode ALPI choice of 75% computer control would correspond to LOA 4 in the task which was developed for Experiment 2, in which the machine had the default response and was 90% accurate, but the human operator could override it when a machine error was suspected. It is thus anticipated that LOA 4 will be a preferred LOA in the subsequent study; however, individual difference measures may not correlate with this choice as they did not correlate with the preferred ALPI choice in Experiment 1.

Experiment 2 was thus conducted to compare automation preferences as influenced by AC level using a target detection decision aid with flexible LOAs, as well as automation preferences on the ALPI in Experiment 1.

CHAPTER 5: LOW AND HIGH LEVELS OF ATTENTIONAL CONTROL IN ADAPTIVE LEVELS OF AUTOMATION

In Experiment 2, each participant undertook the role of the supervisor of a machine decision-aid which had 90% accuracy in correctly identifying targets and nontargets in a series of perceptual tasks. Participants with low and high attentional control scores were recruited. They were given the choice of four different LOAs in which to complete the perceptual tasks, based upon both their preferences for operator-machine interaction and their needs in order to assure accurate target identification. Thus, both LOA preferences and task performance were assessed, along with reports of subjective workload.

Hypotheses

1. Operators with low attentional control will have worse performance in dual-task conditions in an automated system due to the increased number of channels to be monitored. This is largely driven by impairments in engaging, focusing, and disengaging from information when appropriate.
    a. However, this dual-task performance decrement may be attenuated if they select higher LOAs to aid them.
    b. Because of the higher task demand in dual-task conditions, they will select higher LOAs. Higher LOAs will be preferred in general.
    c. As a result of selecting higher LOAs for assistance, the dual-task performance decrement may be attenuated.
    d. Workload reports will be higher under dual-task conditions and at lower LOAs. There will be an interaction between LOA and task demand, with

higher workload reported for dual-tasking conditions executed at lower LOAs. If higher LOAs are selected, the reported workload will be reduced.

2. Operators with high attentional control have greater ability in engaging, focusing, and disengaging their attention when necessary; this generally renders them more adaptive to various levels of task demand. This will enable them to complete dual-task conditions more successfully compared to those with lower attentional control; however, they will likely also perform single-task conditions just as well if not better than those with low attentional control.

    e. The ability to perform tasks successfully will enable those with high attention control to chose lower LOAs in both single- and dual-task conditions, however, their adaptability will give them the flexibility to select any LOA. They will prefer lower LOAs in general.
    f. They will report lower overall workload across conditions.

Methodology

Experimental Participants

Thirty-six participants (28 females, eight males) were recruited from the original subject pool assessed in Experiment 1. They ranged in age from 18 to 34 (*M* = 22.25, *SD* = 3.524). Recruitment criteria included consent to be contacted for recruitment to subsequent studies (see Appendix F) and Attentional Control scores in the lower and upper quartile ranges as described in the Experiment 1 methodology (see Figure 9 for a histogram depicting the low and high AC ranges). Individual difference measures collected in Experiment 1 were retained for analyses in Experiment 2. Participants were compensated financially at the rate of $7.50/hour and also received course credit for completing the session.

Of the participants from the Experiment 1 subject pool who returned for Experiment 2, the AC scores for the ‘Low AC’ group ranged from 36 to 44 (*M* = 39.94, *SD* = 2.461). The low AC group consisted of 18 participants: four males and 14 females. Among the low AC male participants, the mean age was 21.75 years (*SD* = 1.708) and the mean AC score was 40.50 (*SD* = 1.732). Among the low AC female participants, the mean age was 22.57 years (*SD* = 3.631) and the mean AC score was 39.79 (*SD* = 2.665).

The AC scores for the ‘High AC’ group ranged from 56 to 73 (*M* = 60.22, *SD* = 4.209). They affirmed that they had normal-to-corrected vision and do not have known hearing impairments. The high AC group consisted of 18 participants: four males and 14 females. Among the high AC male participants, the mean age was 21.25 years (*SD* = 1.893) and the mean AC score was 61.75 (*SD* = 2.754). Among the low AC female participants, the mean age was 22.36 years (*SD* = 4.272) and the mean AC score was 59.79 (*SD* = 4.526).

There were 135 participants who did not agree to return for Experiment 2. Of these participants, there were 42 were males with AC scores ranging from 37 to 69 (*M* = 52.5, *SD* = 8.25). There were 93 females with AC scores ranging from 27 to 78 (*M* = 52.18, *SD* = 8.83).

The ratio of female to male participants was roughly representative of the population of undergraduate psychology majors who were eligible to participate in the study. Previous literature has indicated that sex differences in signal detectability have been not been found in temporally-oriented perceptual monitoring tasks (Dittmar, Warm, Dember, & Ricks, 2001). Furthermore, current automation systems are not designed differently for male and female operators; thus, until future studies illuminate any possible gender differences in automation use, the current study is based upon the presumption that both genders interact with technology similarly.

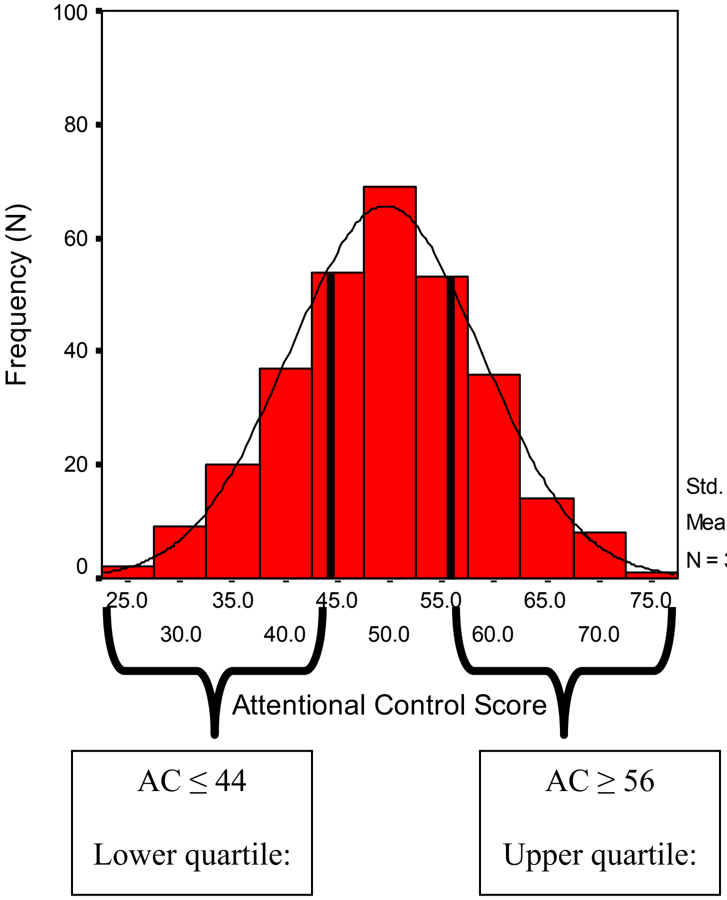


Figure 8. Histogram which features the ‘Low AC’ and ‘High AC’ overlay upon the entire AC distribution. Participants were selected from the high and low AC groups.

Experimental Apparatus

All automation tasks were conducted on a Pentium 4 PC, and the display was presented on a standard 15.5-inch CRT. Sound was delivered through Sony noise-canceling headphones. The automation tasks were created using LabView 7.1 development software.

Experimental Materials

The NASA-Task Load Index (NASA-TLX: Hart & Staveland, 1988; see Appendix G) is a self-reported, subjective workload assessment, which uses six independent rating scales to derive measures of frustration, performance, temporal demand, physical demand, effort, and mental demand. The summation of these six subscales produces an overall assessment of the workload experienced during the task. Task instructions were presented in a booklet and read aloud by the experimenter as the participant read along (see Appendix H insert task instructions).

Experimental Design

A 4 (selectable LOA: Fully Manual, Mostly Manual, Mostly Machine, Equal Manual-Machine Allocation) by 2 (Modality: Auditory, Visual) by 2 (Load: Single-Task, Dual-Task) within-subjects experimental design was used. AC group (high and low) was a between-subjects variable.

The Automation Program

Participants completed audiovisual perceptual tasks and were instructed to select a level of computer decision-making assistance using an automated system interface. The visual target detection task was presented on the PC monitor, and the auditory target detection task was delivered through headphones. Lines appeared on the left and right side of the screen separately, and tones were presented in the left and right earphones separately. This generated four types of stimuli which were presented singularly and randomly, but never simultaneously. Participants identified each stimulus type after it was presented by using a mouse to click on the buttons on the display labeled ‘Target’ and ‘Non’. The term ‘non’ was used instead of ‘nontarget’ to increase saliency of each term and therefore make them less confusable with each other during the tasks. See Figure 10 for an image of the interface.

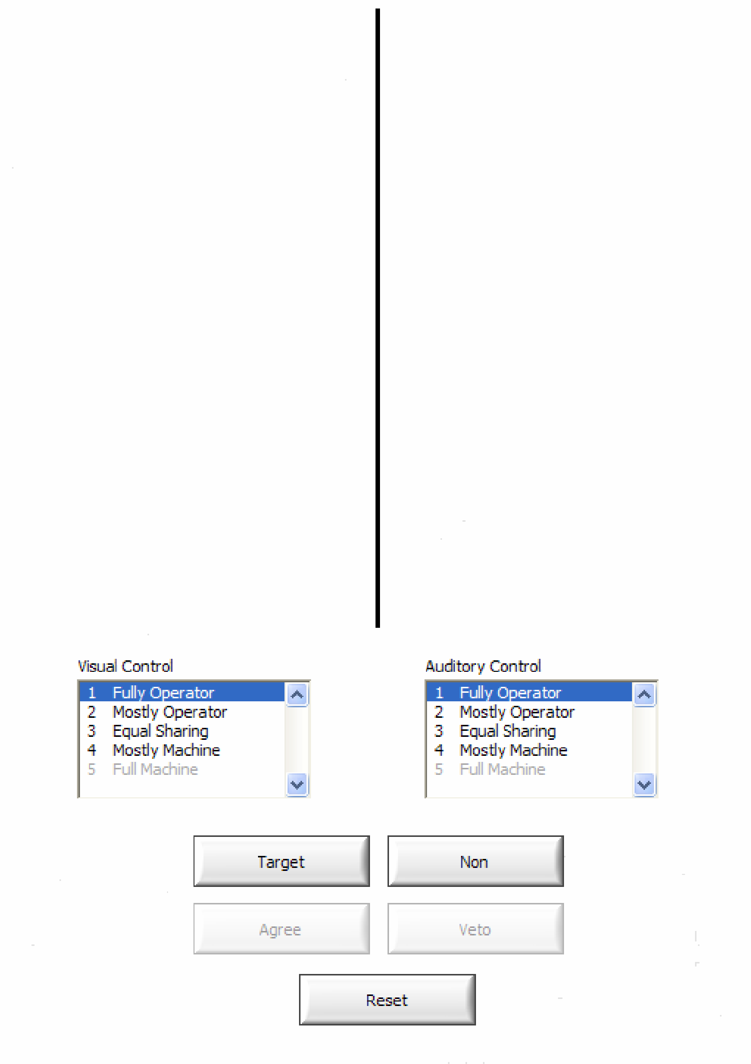


Figure 9. Automation program interface example.

The vertical black division line determines the left and right sides of the screen. Five LOAs are listed in both the Visual Control and Auditory Control mode menus. This example features the audiovisual combined condition. The current LOA is highlighted in blue; both the Visual and Auditory Control modes are set at ‘1 Fully Operator’. “5 Full Machine” is not an option for either task, so the option on the menu has a gray cast to indicate that the option is disabled.

The automation system was 90% reliable in terms of target detection accuracy. Participants were informed of this imperfect reliability and that therefore fully relying on automation may result in false alarms and missed targets. The automation application was programmed to operate at a reliability of 90% in terms of both target and non detection answers in the mostly machine LOA and suggestions made to the operator in the mostly manual LOA, as well as the intermediary LOA in which both suggestions and answers were provided. This 90% reliability yielded 18 hits, 18 correct rejections, two false alarms, and two misses per task.

Visual-only target detection task. In the visual-only target detection task, participants differentiated between black vertical lines (32mm in height) presented at either a short (125 ms) or long (250 ms) period of time upon a white background. Targets were lines appearing for the shorter period of time, while nons were lines appearing for the longer period of time. Target and non lines as well as target and non tones were presented in random orders.

During the visual-only task, the auditory task was fully executed by the computer, and the participant was instructed to ignore the auditory task, which consisted of the tones and the machine responses to the tones. However, they were instructed to monitor the machine's functionality by pressing a 'Reset' button on the screen using a mouse in the event that the machine failed to provide identifications of the auditory stimuli. These failures were described as making too many mistakes (in the event that the participant was able to monitor the machine's detections) and/or cessation of identifications altogether. Unbeknownst to the participants, machine failures were not programmed to occur. The 'Reset' function was implemented to ensure the participant maintained awareness of the machine functionality even while fully automated, and this system monitoring replicates the supervisory role that many human operators have in human-automation interaction. It was also intended to ensure that participants wore the headphones during the visual task and were exposed to the auditory stimuli even when instructed to ignore them.

Auditory-only target detection task. In the auditory-only target detection task, participants distinguished between sounds (400 Hz C4 triangle waveform; see Droit-Volet, 2004) that were either short (200 ms) or long (250 ms) in presentation time. During the auditory-only task, the visual detection task was fully executed by the computer, and the participant was asked to ignore the visual task, which consisted of the lines and the machine responses to the visual stimuli. As in the visual task, they were instructed to click on the 'Reset' button in the event of an automation failure in the visual task.

Audiovisual combined target detection task. In the audiovisual combined task, both the audio and visual target detection tasks were presented simultaneously, and the participant and machine were responsible for performing none, one, or both tasks according to the participant's chosen LOA and method of operating the LOA. As in the visual-only and auditory-only tasks, both lines and tones were presented in random orders, but did not overlap temporally or spatially such that only one stimulus was presented at a time.

A temporal discrimination task was chosen because stimulus presentation time is a characteristic that can be manipulated in both auditory and visual displays. These visual and auditory short and long temporal values have been psychophysically equated for discriminability (see Szalma et al., 2004). The difference between the short and long stimuli was greater in the visual task than in the auditory task (i.e., 125 ms vs. 50 ms delta, respectively); this is to compensate for the finding that temporal perception in the auditory modality is more accurate than in the visual modality (Dember & Warm, 1979).

Before the first training session, each participant matched the apparent loudness of the auditory stimuli was to the apparent brightness of the visual stimuli using a cross-modality matching technique (see Gescheider, 1997). This matching was intended to psychophysically equate the perceptual intensity of both visual and auditory stimuli, thereby precluding a bias resulting from greater saliency of stimuli of either modality.

LOA Algorithms. Each of the five different LOAs followed specific algorithms of operator-machine response patterns. A brief outline of the different LOAs is provided below; detailed descriptions of the functionalities of each LOA are available in the participants' instruction booklet (see Appendix H). The participant, who was referred to as the 'operator' in the task, was instructed to respond to stimuli that were his/her responsibility (i.e., all stimuli in '1 Fully Operator', all stimuli in '2 Mostly Operator' and stimuli on left side of the screen and in the left ear in '3 Equal Sharing') by using a mouse to click on buttons labeled 'Target' and 'Non' on the display. The participant was also given the option to veto or agree with the machine's decisions for stimuli that were its responsibility (i.e., all stimuli in '4 Mostly Machine' and stimuli on the right side of screen and in the right ear in '3 Equal Sharing') by using a mouse to click on buttons labeled 'Agree' and 'Veto' (see Figure 10 for an image of the display). Because two stimuli never occurred at the same time, it was possible to use the same set of buttons to respond to visual and auditory stimuli, and also to separate and analyze the participants' responses made to the visual and auditory stimuli in the combined condition.

The '1 Fully Operator' condition corresponds to the LOA of 1 of the ALPI (administered in Experiment 1; see Appendix D), while '2 Mostly Operator' condition corresponds to LOA 2, the '3 Equal Sharing' condition corresponds to LOA 3, the '4 Mostly Machine' condition corresponds to LOA 4, and the '5 Full Machine' condition corresponds to LOA 5.

5 Full Machine Algorithm. The machine identified stimuli on both sides of the screen in the visual task, or in both ears in the auditory task. The operator did not have the opportunity to veto or agree with the machine's identifications. The operator was not required to attend to the lines or tones, but was still responsible for ensuring that the machine decision aid was still functioning by looking for its answer messages (i.e., 'Target' and 'Non' dialog boxes after each line) or listening for its voice messages in the earphones (i.e., 'Target' and 'Non' voices after each tone). If the operator suspected an automation failure, s/he was instructed to click on the 'Reset' button. The steps were as follows:

1. Stimulus presented.
2. Machine identified the stimulus with 90% reliability.
    - Visual Task: If the machine identified it as a target, it displayed a dialog box reading, "Target". If it identified it as a non, then it displayed a dialog box reading, "Non".

- Auditory task: If the machine identified it as a target, then it delivered a voice that said, "Target". If the machine identified it as a non, then it delivered a voice that said, "Non".

3. Next stimulus appeared.

4 Mostly Machine Algorithm. The machine identified stimuli on both sides of the screen in the visual task, or in both ears in the auditory task. The operator was given a three-second response time window to veto the machine's decision if s/he disagreed with the machine's decision by clicking on 'Veto'. The operator could also indicate agreement with the machine responses by clicking on 'Agree', although not clicking 'Agree' was equivalent to agreeing with the machine, as the machine answer was the default. The operator was not required to attend to or respond to the machine's decisions; rather, s/he had the option to monitor the machine's responses. In the cases that the operator: a) did not veto any of the machine responses, or b) agreed with all of them (i.e., all of the machine's answers were processed) the machine automatically identified 90% of the stimuli correctly and the participant's final score (measured in proportion correct) for that task would therefore be 90%. Operators were instructed to supervise the machine if they wished to correct any of the machine's incorrect decisions, with the goal of attaining more than 90% correctness. The steps were as follows:

1. Stimulus presented.
2. Machine identified the stimulus with 90% reliability.
    - Visual Task: If the machine identified it as a target, it displayed a dialog box reading, "Target". If it identified it as a non, then it displayed a dialog box reading, "Non. Veto?"

- Auditory task: If the machine identified it as a target, then it delivered a voice that said, “Target. Veto?” If the machine identified it as a non, then it delivered a voice that said, “Non. Veto?”

3. The operator had the option to veto the machine’s decision by clicking on the ‘Veto’ button to correct the machine’s decision, or ‘Agree’ to affirm the decision within the three-second response window.
4. Next stimulus appeared.

3 Equal Sharing Algorithm (half operator control, half machine control).

The system used the ‘2 Mostly Operator’ LOA (described in further detail below) for lines appearing on the left side of screen and tones sounding in the left earphone, and the ‘4 Mostly Machine’ for lines appearing on the right side of screen and tones sounding in the right earphone.

- Left field: The operator was responsible for identifying stimuli on the left side of the screen and in the left earphone. The operator was required to identify each stimulus as a target or non first, and then the machine provided a suggestion in the form of its own independent identification of the stimulus (i.e., a dialog box in the visual task or a verbal message in the auditory task). The operator then had the option to use the machine’s suggestion as his or her answer instead by clicking on ‘Agree’ after the machine provided its suggestion. The operator had a three-second response window to respond to the machine suggestion.
- Right field: The machine identified stimuli on the right side of the screen and in the right earphone. The machine identified each stimulus in the right fields first

as a target or non, and then the operator had the option to attend to the stimuli in the right fields and veto these decision if s/he wished to do so.

2 Mostly Operator-Controlled Algorithm. The operator identified stimuli on both sides of the screen in the visual task, or in both ears in the auditory task first. The machine then made an independent suggestion of whether the stimulus was a target or non, which was 90% reliable. The operator then had the option to change his or her decision based on the machine's suggestion by clicking on 'Agree' or 'Veto'.

1. Stimulus presented.
2. Visual and/or Auditory Task: Operator identified it as a target or non by clicking on the 'Target' or 'Non' buttons, respectively.
3. Machine offered a suggestion as to whether the stimulus was a target or non with 90% reliability.
    - Visual Task: If the machine identified it as a target, then it displayed a dialog box that reads, "Target. Veto?" If it identified it as a nontarget, then it displayed a dialog box reading, "Non. Veto?"
    - Auditory task: If the machine identifies it as a target, then it delivered a voice that said, "Target. Veto?" If the machine identifies it as a nontarget, then it delivers a voice that says, "Non. Veto?"
4. The operator had the option to change his/her answer based on the machine's suggestion by clicking on 'Agree' to affirm the machine suggestion, or 'Veto' to disagree with the machine suggestion within in three-second response window.
5. Next stimulus appeared.

1 Fully Operator-Controlled Algorithm. The operator identified stimuli on both sides of the screen in the visual task, or in both ears in the auditory task. The machine did not make any suggestions to the operator regarding whether a stimulus is a target or non.

1. Stimulus presented.
2. Visual and/or Auditory Task: Operator identifies it as a target by clicking on the 'Target' button or as a nontarget by clicking on the 'Non' button.
3. Next stimulus appears.

In the single task visual-only condition, participants selected the LOA they preferred for only the visual task, while the auditory task was fully automated and fixed in the "5 Full Machine" setting. Conversely, in the single task auditory-only condition, participants selected the LOA they preferred for only the auditory task; the visual task was be fully automated and fixed in the "5 Full Machine" setting. In the dual-task audiovisual combined task, participants selected the LOA they preferred for both the visual and auditory target detection tasks.

Participants could select among four of the five LOAs: '1 Fully Operator', '2 Mostly Operator', '3 Equal Sharing', and '4 Mostly Machine'. They engaged the LOA of choice by using the mouse to click on the LOA from a menu on the screen. The Visual Control Mode referred to the level of machine assistance for identifying the line stimuli, while the Auditory Control mode referred to the level of machine assistance for identifying the tones. Participants changed the LOAs by clicking on from the control mode menus on the screen as times as they wish at any time during all three tasks. Upon selecting an LOA, the machine responded to this request and then allocated task control accordingly. See Figure 10 above for an image of the display.

The ‘5 Full Machine’ setting was only engaged for the task which was automated during the single-task conditions. It could never be selected by the participant in any task lest the operator be free of any detection responsibilities. During the single-task conditions, the ‘5 Full Machine’ option on the menu appeared highlighted for the task which was fully automated, and all the other LOA options for that task appeared disabled and gray-cast. Also during the single task condition, the ‘5 Full Machine’ option on the menu was disabled and gray-cast to indicate it was not selectable in the task which was being completed by the participant. In the dual-task condition, the ‘5 Full Machine’ was not be an option for either task and the participant was not be able to select it. Therefore, it appeared disabled for both tasks. When ‘5 Full Machine’ was engaged during the single-task condition, participants were instructed to ignore the stimuli it was identifying, but to monitor the automation (i.e., listen for its voice in the auditory task and looking for its dialog box in the visual condition) to ensure that it is still functioning and clicking on ‘Reset’ when dysfunctions occurred. See Table 4 for a synopsis of the possible LOAs to be chosen by the participants for each of the three tasks.

Table 4. Possible LOA choices to be made by the participant during the three tasks.

	Visual-Only Task		Auditory-Only Task		Audiovisual Combined Task	
LOA	Left Screen	Right Screen	Left Earphone	Right Earphone	Left Screen/ Earphone	Right Screen/ Earphone
5:Full Machine	Fixed for Audio Task	Fixed for Audio Task	Fixed for Visual Task	Fixed for Visual Task	Not an option	Not an option
4:Mostly Machine	Machine responds to stimuli. Operator has option to override machine responses.	Machine responds to stimuli. Operator has option to override machine responses.	Machine responds to stimuli. Operator has option to override machine responses.	Machine responds to stimuli. Operator has option to override machine responses.	Machine responds to stimuli. Operator has option to override machine responses.	Machine responds to stimuli. Operator has option to override machine responses.
3: Equal Sharing	Operator decisions required. Operator has option to change his original response based on machine's suggestion	Machine responds to stimuli. Operator has option to override machine responses.	Operator decisions required. Operator has option to change his original response based on machine's suggestion	Machine responds to stimuli. Operator has option to override machine responses.	Operator decisions required. Operator has option to change his original response based on machine's suggestion	Machine responds to stimuli. Operator has option to override machine responses.
2: Mostly Operator	Operator decisions required. Operator has option to change his original response based on machine's suggestion.	Operator decisions required. Operator has option to change his original response based on machine's suggestion.	Operator decisions required. Operator has option to change his original response based on machine's suggestion.	Operator decisions required. Operator has option to change his original response based on machine's suggestion.	Operator decisions required. Operator has option to change his original response based on machine's suggestion.	Operator decisions required. Operator has option to change his original response based on machine's suggestion.
1: Full Operator	Operator decisions required.	Operator decisions required.	Operator decisions required.	Operator decisions required.	Operator decisions required.	Operator decisions required.

Experimental Procedure

Participants were presented with informed consent forms that affirms they were at least 18 years of age and explained their rights as participants, including the right to withdraw their participation at any time without penalty (see Appendix H).

In the training session, task instructions were presented in a booklet and read aloud by the experimenter as the participant read along (see Appendix H). These instructions explained the nature of the tasks, demonstrations of target and non lines and tones, the machine decision-aid system, and the five different LOAs. These instructions were also punctuated with demonstrations of each LOA in use and practice with the audiovisual perception tasks for two minutes at each LOA, beginning with LOA 5, and then descending the levels until LOA 1 was reviewed. The duration of the training period was determined by the results of pilot testing which indicated that participants felt comfortable using the various LOAs and were also capable of completing the perceptual tasks with approximately 90% accuracy. The next LOA was discussed once all questions for the current LOA were answered and participants affirmed they were familiar with using that LOA. This training allowed participants to become familiarized with the automation interface. According to Parush and Auerbach (2005), there may be a temporary, yet significant performance cost associated with both user- and automation-controlled changes in system adaptation. The training session thus served as a learning period which provided the participants with confidence and ability to make adaptation changes and accept decisions proposed by the system in the actual experimental task. An upright chart remained on the desk next to the computer monitor which summarized each of the five LOAs to serve as a reminder regarding how each level functioned (see Appendix J).

After the training sessions, participants began the experimental sessions, consisting of the same three tasks that they completed in the practice session: auditory-only (single-task), visual-only (single-task), and the audiovisual combined (dual-task).

The order of the single-task conditions were counterbalanced across participants, while the audiovisual combined task was completed last. The instruction set indicated to the participants that the participant with the best overall performance across all three tasks would receive a $100 cash reward.

A default LOA was initially set by the experimenter; this was the LOA at which the task began, and participants could change the default LOA as soon after the task began to a different LOA of choice. Half of the participants (N = 18) began the tasks at a default LOA 1, while the other half began the task at a default LOA 4.

Each perceptual task consisted of 80 events (40 visual and 40 auditory) and lasted approximately eight minutes. In the single-task visual-only condition, participants were still exposed to the tones, but instructed to ignore them; thus only the 40 visual trials comprised their score. Conversely, in the single-task auditory-only condition, participants were still exposed to the lines, but instructed to ignore them; only the 40 auditory trials comprised their score. In the audiovisual combined task, participants were responsible for identifying both lines and tones; thus all 80 trials comprised their score. Further, each event type was divided between the left/right screens and left/right earphones. For the single-task conditions, the event rate averaged at 5 stimuli per minute. In the dual-task condition, the event rate doubled to 10 events per minute. The inter-stimulus interval varied between 1500 to 3000 ms. This event rate was chosen based upon pilot testing which indicated approximately 90% detection accuracy, and the tendency for participants to invoke the higher LOAs available to them rather than complete all tasks at '1 Full Operator'. This participant behavior during pilot testing suggested that the automation was useful to the participants and was thus serving its intended purpose.

This particular event rate also mathematically accommodates a 90% signal detection reliability of the machine's responses, such that integer values of discreet hits, misses, false alarms, and correct rejection events can be made. It also ensured that stimuli and response time windows between subsequent events did not overlap, so that participant responses could be linked to the appropriate event or machine suggestion. Finally, it allowed for an equal number of hits, misses, false alarms, and correct rejections for each side of the screen in the visual task and each earphone in the auditory task.

In all LOAs wherein the machine provided answers or suggestions to the operator ('2 Mostly Operator', '3 Equal Sharing', '4 Mostly Machine', and '5 Full Machine'), the 90% machine reliability was programmed as follows: within the 40 events for each modality, the machine committed 18 hits, 18 correct rejections, two misses, and two false alarms. Thus in the dual-task condition, the participant was exposed to 36 hits, 36 correct rejections, four misses, and four false alarms made by the machine.

After each of the three tasks, participants completed the abbreviated version of the NASA-TLX for workload assessment.

After participation was complete, participants were debriefed (see Appendix K).

Results

In Experiment 2, participants were able to change the level of automation (LOA) according to their preference throughout the course of the task. There were four **LOAs** available to the participants: '1 Full Operator', '2 Mostly Operator', '3 Equal Sharing, and ‘4 Mostly Machine’, as described in the methodology. In LOA ‘1 Full Operator’ work was conducted solely by the operator without the machine aid. LOAs 2, 3, and 4 provided the machine decision aid. In LOA 2, the operator responded first and then received a machine suggestion. In LOA 3, the operator responded first to all stimuli in the left field and then received a machine suggestion, while in the right field, the machine responded first and the operator could veto the machine decision. In LOA 4, the machine responded first and then the operator had the option to input her own response if it differed from the machine response. Participant’s preferences for the different LOAs were assessed by examining the percentage of time that participants chose to be in each one of the LOAs, relative to the entire duration of the task.

Four separate tasks were analyzed: visual-only and auditory-only, both of which were the single-task conditions; and the visual portion of the audiovisual-combined task (referred to as ‘visual-combined’) and the auditory portion of the audiovisual combined task (referred to as ‘auditory-combined’), both of which were the dual-task conditions. Thus, two **load** conditions were assessed. Single-task conditions encompass the visual-only and visual-combined tasks, and dual-task conditions encompass the visual-combined and auditory-combined tasks (i.e., the entire audiovisual combined task). Two modalities were differentiated: visual and auditory. **Modality** effects were reported across their respective tasks, such that visual modality effects encompass both the visual-only and visual-combined tasks, while the auditory modality effects encompass both the auditory-only and auditory-combined tasks.

Furthermore, we analyzed the responses of the operator in relation to the decision aid (without and with **machine** aid) hence there were two levels of responses: operator-alone and machine-aided. Operator-alone involved the participants' responses independent of the machine aid, while machine-aided involved the participants' final responses after using the machine decision-aid.

Responses were analyzed using signal detection theory (SDT: Green & Swets, 1966) measures, d' (a measure of the participants' sensitivity to the distinction between targets and nons) and β (beta: a measure of the participants' response bias towards conservatism or leniency in reporting stimuli as targets). See Macmillian and Creelman (2005) for a complete review of SDT.

Preferences for LOAs

A 2 (Modality) x 2 (Load) x 4 (LOA) within-subjects repeated measures ANOVA was conducted. AC was a between-subjects variable.

See the ANOVA Table (Table 5) for a summary of the main effects and interactions for LOA preference in terms of proportion of time spent in each LOA.

Table 5. Analysis of Variance for LOA Preference (Time in LOA)

Source	*df*	*F*	η^2	*p*
		Between subjects		
Attentional Control (AC)	1	.838	.024	.366
S within-group error	34	(.007)		
		Within subjects		
Modality	1	.260	.008	.613
Modality x AC	1	.324	.009	.573
Modality x Load	1	.113	.003	.739
Modality x Load x AC	1	.517	.015	.477
Modality x S within group error	34	(.007)		
Load	1	3.879	.102	.057
Load x AC	1	.572	.017	.454
Load x S within group error	34	(.006)		
LOA	2	1.429	.040	.247
LOA x AC	2	1.233	.035	.298
LOA x Modality	2	10.344**	.233	.000
LOA x AC x Modality	2	2.577	.070	.083
LOA x Load	1.588	3.218*	.086	.046
LOA x Modality x Load	2	.754	.022	.474
LOA x Load x AC	2	.138	.004	.871
LOA x Load x Modality x AC	2	.071	.002	.932
LOA x S within group error	68	(.876)		

Note. Values enclosed in parentheses represent mean square errors. *S* = subjects.

**p* < .05. ***p* < .01.

There was a significant main effect for LOA, $F(3, 105) = 6.841$, $p < .0001$. Post hoc tests conducted with Tukey's HSD ($q = 3.825$, $t = 2.696$) revealed that that significantly more time was spent at levels '1 Fully Operator' than '3 Equal Sharing', $t(35) = 3.257$, $p < .005$. Significantly more time was spent at '2 Mostly Operator' than '3 Equal Sharing', $t(35) = 5.263$, $p < .0001$. Significantly more time was also spent at '4 Mostly Operator' than '3 Equal Sharing', $t(35) = -4.862$, $p < .0001$.

Sphericity violation indicated by a significant Mauchly's Test warranted an adjusted *df* value. The Greenhouse and Geiser (1958) estimate was selected based on a rationale described by Girden (1992; see also Field, 2000).

The mean proportion of time spent at LOA '1 Full Operator' was .218 (*SD* = .346), .397 at '2 Mostly Operator' (*SD* = .420), .024 (*SD* = .070) at '3 Equal Sharing', and .361 (*SD* = .414) at '4 Mostly Machine'. See Figure 11 for a depiction of the time proportions.

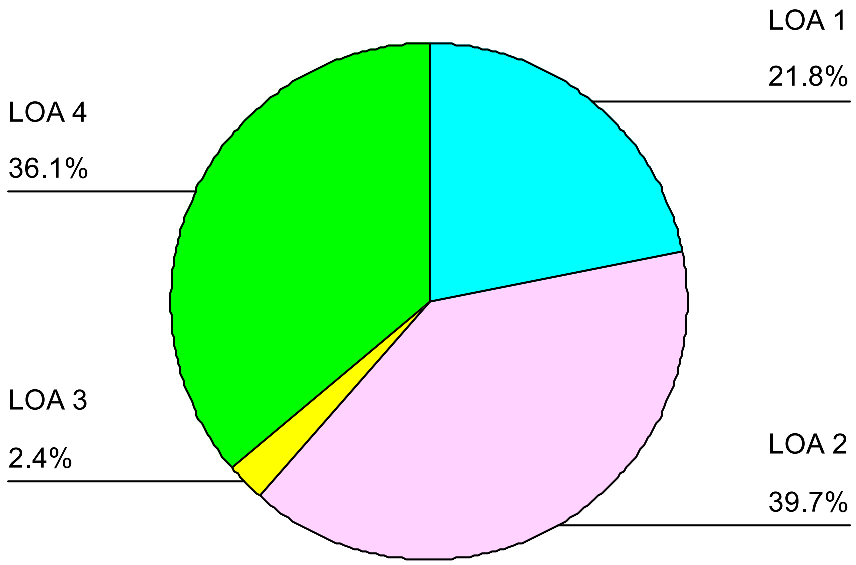


Figure 10. Proportion of time spent at each LOA collapsed across tasks.

There was no significant interaction between the initial LOA which was selected for each participant as the default LOA and the proportion of time spent at each LOA, $F(2, 66) = .368$, $p > .05$. This indicates that the once the task began, participants changed their LOA and they were not necessarily anchored to the system's default level.

The main effect for LOA was predominantly due to the fact that LOA '3 Equal Sharing' was very infrequently chosen when compared to the other LOAs. This discrepancy led to a significant Mauchley's test of sphericity for LOA and all interactions among it and other variables (e.g., modality *LOA, load*LOA, and modality*load*LOA), all $p < .005$. This was thought to mask effects seen with the other three LOAs; as such, LOA '3 Equal Sharing' was removed from the ANOVA and levels

1, 2, and 4 were assessed for significant effects in a revised repeated-measures ANOVA. All further analyses were conducted by considering only LOAs 1, 2, and 4.

The revised experimental design was a 2 (Modality) x 2 (Load) x 3 (LOA) repeated measures ANOVA.

There was not a significant main effect for LOA, *F*(2, 68) = 1.429, $p > .05$. Proportion of time spent at LOA 1 (*M* = .221, *SD* = .346), LOA 2 (*M* = .402, *SD* = .420), and LOA 4 (*M* = .352, *SD* = .414) were thus considered to be statistically the same.

There was a significant interaction between modality and LOA, *F*(2, 68) = 10.344, $p<.0001$. Post hoc paired samples t-tests showed that in the auditory tasks, operators spent more time at LOA 2 than they did in the visual tasks (*M*=.453, *SD*=.440, and *M* = .351, *SD* = .439, respectively), t(35) = 2.390, $p < .05$. Adjustments made by Tukey's HSD (q = 4.268, t = 3.018), however, did not render this difference significant. Cohen's *d* was .229.

In the visual modality, the total time at LOA 4 was significantly higher than the time spent at LOA 2 (*M* = .439, *SD* = .425, and *M* = .192, *SD* = .346, respectively), *t*(35) = 2.307, $p < .05$. Adjustments made by Tukey's HSD (*q* = 4.268, *t* = 3.018), however, did not render this difference significant. Cohen's *d* was .822. In the visual tasks, LOA 4 was used more than it was in the auditory tasks (*M* =. 439, *SD* = .425, and *M* = .293, *SD* = .462, respectively), *t*(35) = 2.743, $p = .01$. However, adjustments made by Tukey's HSD (*q* = 4.268, *t* = 3.018) did not render this difference significant. Cohen's *d* was .485. See Figures 12 and 13 for depictions of proportion of time spent at each LOA in the visual modality and in the auditory modality, respectively.

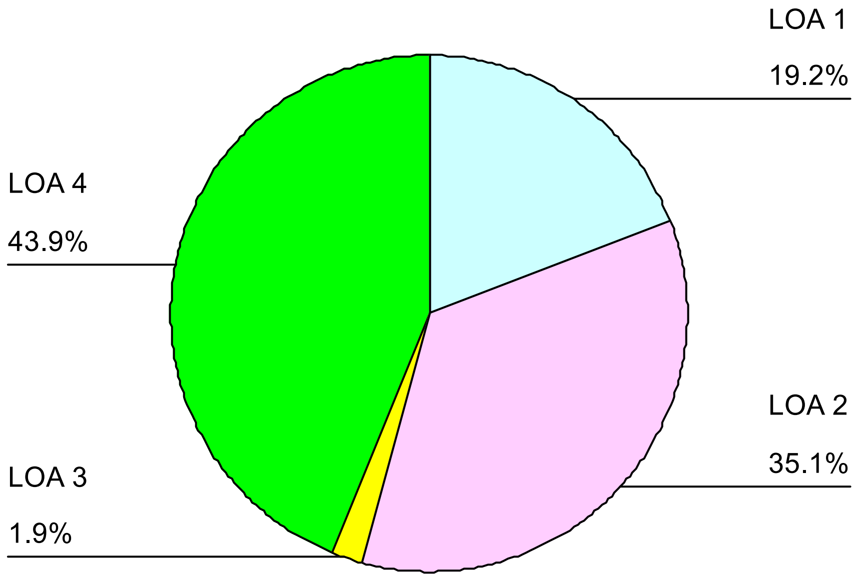


Figure 11. Proportion of time spent at each LOA in the visual modality.

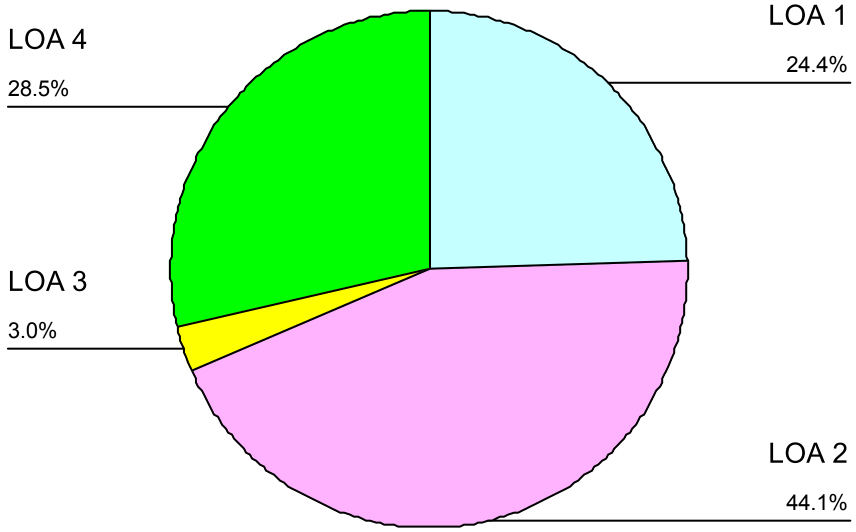


Figure 12. Proportion of time spent at each LOA in the auditory modality.

There was a significant interaction between LOA and load, *F*(2, 68) = 3.218, $p < .05$, however, Mauchley's test of sphericity was significant and thus the Greenhouse-Geisser adjustment for *df* was used. With this adjustment, the interaction between LOA and load was marginal, *F*(1.588, 54.006) = 3.218, $p = .059$. Post hoc paired-sampled t-tests showed that use of LOA 4 in the dual task conditions (across modality) was significantly higher than in the single task conditions ($M = .405$, $SD = .426$, and $M = .298$, $SD = .388$, respectively), t(35)=2.250, p<.05, $d = .26$. Adjustments made by Tukey's HSD ($q = 4.268$, $t = 3.018$), however, did not render this difference significant.

The interaction among modality, load, and LOA was not significant, *F*(2, 61.819) = .754, $p > .05$, however, a trend was seen in more time spent at LOA 4 in the visual-combined task than in the visual-only task (*M* = .528, *SD* = .44, and *M* = .361, *SD* = .41, *d* = .39). See Figures 14, 15, 16, and 17 for depictions of proportion of time spent at each LOA in the visual-only, auditory-only, visual-combined, and auditory-combined tasks, respectively.

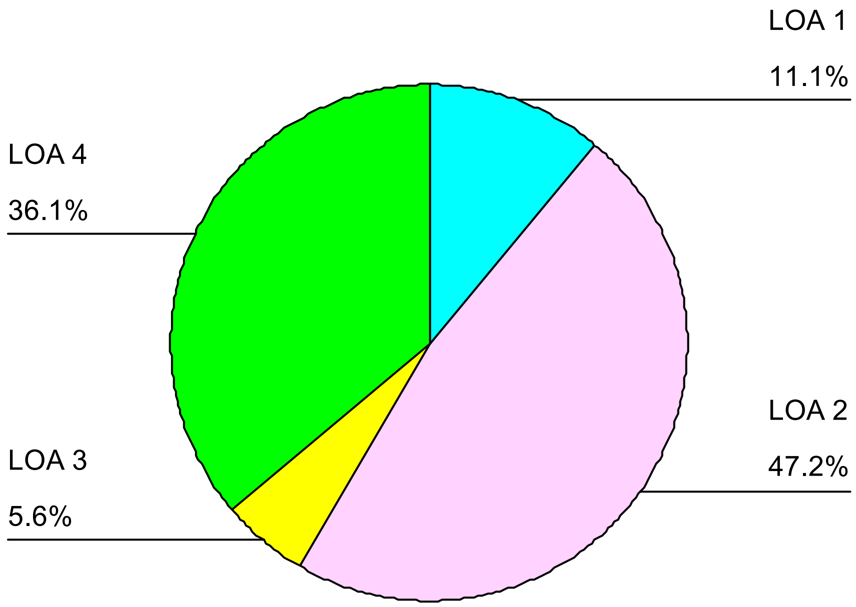


Figure 13. Proportion of time spent at each LOA in the visual-only task.

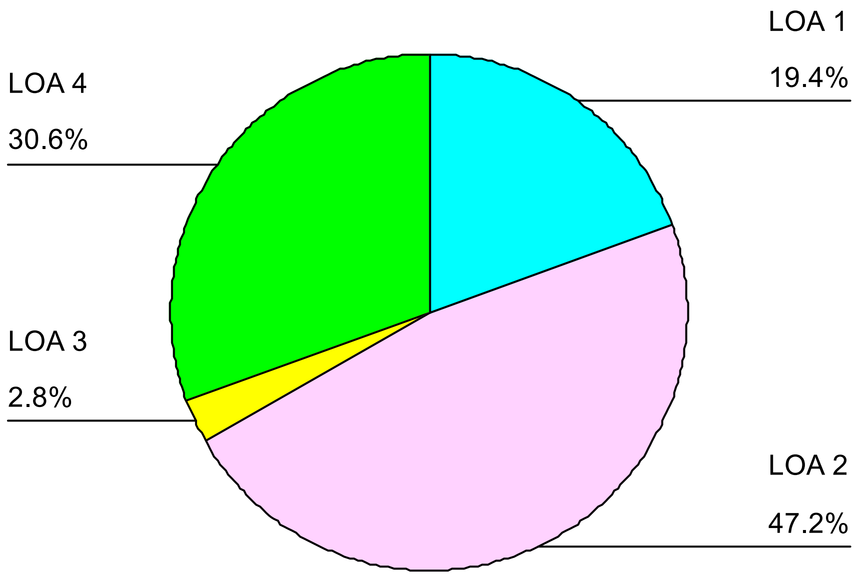


Figure 14. Proportion of time spent at each LOA in the auditory-only task.

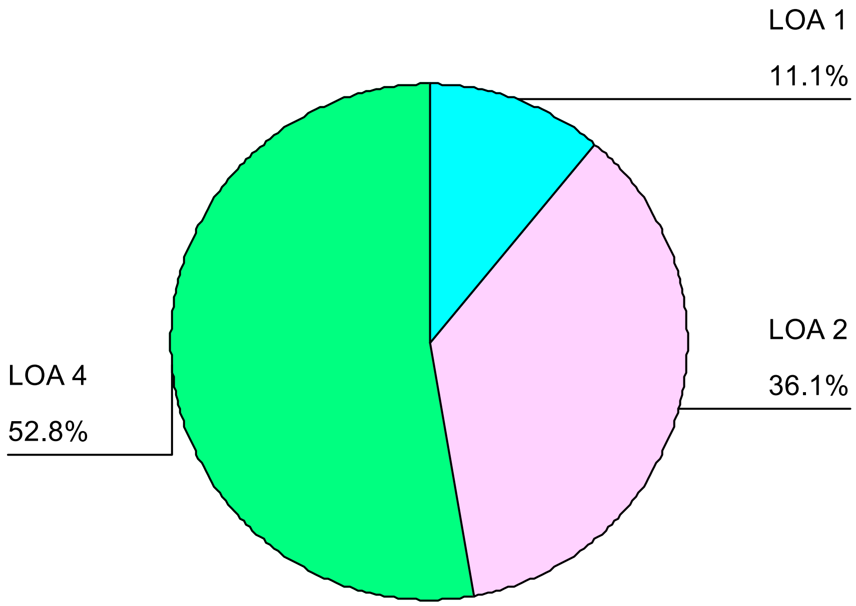


Figure 15. Proportion of time spent at each LOA in the visual-combined task.

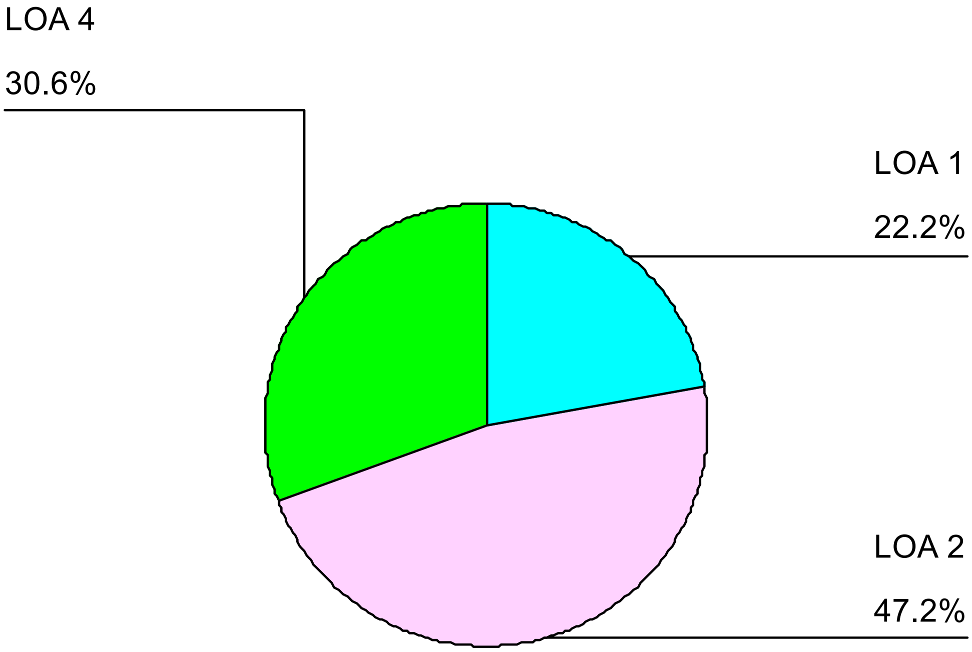


Figure 16. Proportion of time spent at each LOA in the auditory-combined task.

Although there was no significant interaction between LOA and AC, F(2, 68) = .587, p > .05, there was a trend for those with low AC to use higher LOAs than those with high AC across all four tasks (see Figure 18). A priori test of simple effects indicated that those with low AC spent significantly more time at LOA 4 than LOA 1 (M = .482, SD = .467, and M = .096, SD = .234, respectively), t(17) = 2.851, p < .05, *d* = 1.17. There was also a trend for those with low AC individuals to spend more time at LOA 4 than those with high AC (M = .441, SD = .377, and M = .291, SD = .447, respectively, *d* = .56). Conversely, those with high AC spent more time at LOA 1 than those with low AC (M = .304, SD = .419, and M = .138, SD = .237, respectively, *d* = .56). Also, see Figures 19 and 20 for bar graphs depicting the proportion of time spent at each LOA per task, for the low AC group and high AC group, respectively).

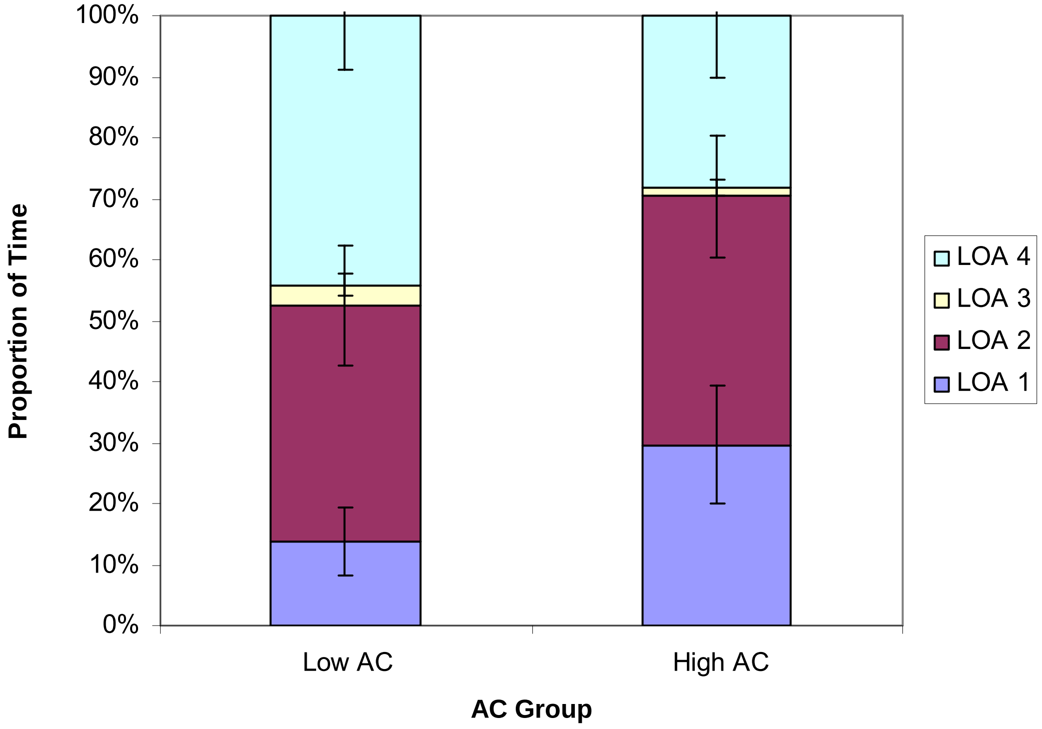


Figure 17. Overall proportion of time spent at each LOA collapsed across tasks for low and high AC groups.

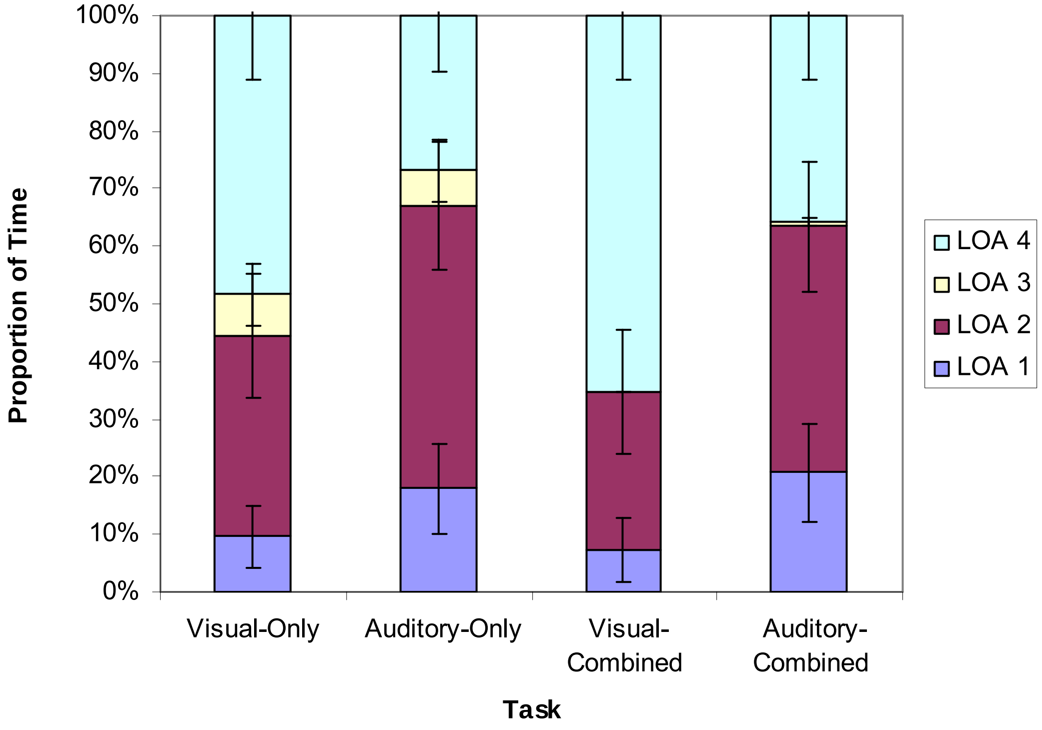


Figure 18. Overall proportion of time spent at each LOA for each task for the low AC group.

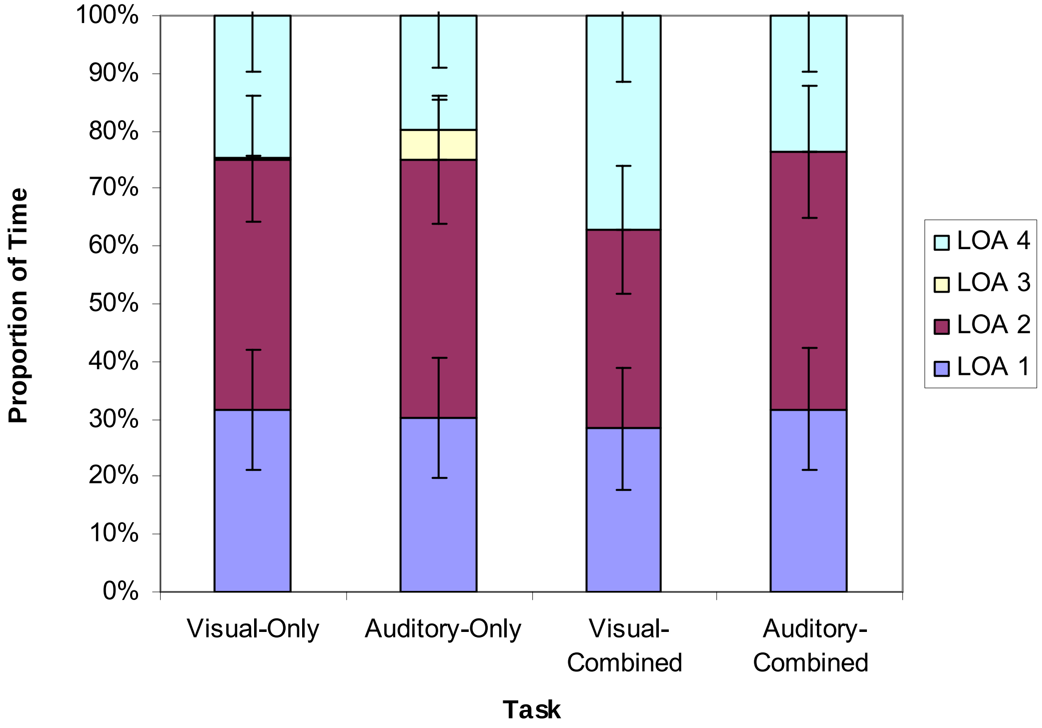


Figure 19. Overall proportion of time spent at each LOA for each task for the high AC group.

Operator-Machine Interaction

Participants were permitted to change LOAs by clicking on them from the menu list as often as they wished during the span of each task until they found a level at which they were comfortable. They were also permitted to change this level as many times as they wished over the course of the task. In the visual-only task, the mean number of LOA changes was 1.14 (*SD* = 1.13), while in the auditory-only task, the mean was .94 (*SD* = .63). In the visual-combined, the mean was .86 (*SD* = .64), and in the auditory-combined task, the mean was 1.14 (*SD* = 1.33).

Although there were no significant effects for AC group, there was a trend for low AC operators to change LOAs more often than those with high AC (see Figure 21). In the visual-only task, the mean number of changes was 1.28 (*SD* = 1.074) for those with low AC, and 1.00 (*SD* = 1.188) for those with high AC, *d* = .37. In the auditory-only task, the mean number of changes was 1.06 (*SD* = .639) for those with low AC, and .83 (*SD* = .618) for those with high AC, *d* = .51. In the visual-combined task, the mean number of changes was .94 (*SD* = .725) for those with low AC, and .78 (*SD* = .647) for those with high AC, *d* = .31. In the auditory-combined task, the mean number of changes was 1.39 (*SD* = 1.650) for those with low AC, and .89 (*SD* = .900) for those with high AC, *d* = .48.

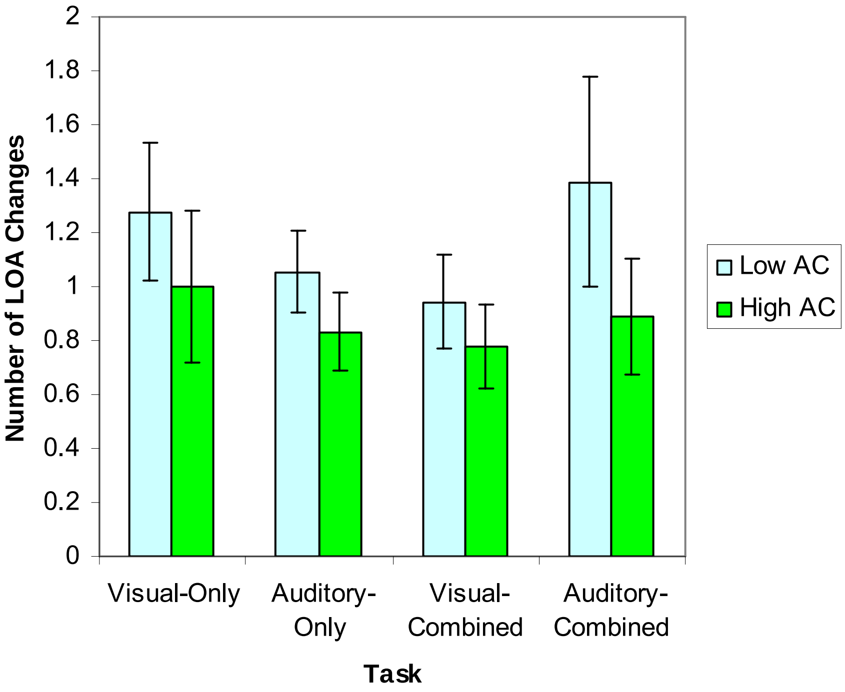


Figure 20. Mean number of LOA changes per task grouped by AC.

Participants opted to agree with the machine response or veto it. Figure 22 depicts the mean number of times 'Agree' or 'Veto' was clicked on during each of the four tasks. The mean number of 'Agree' and 'Veto' responses was significantly higher in the auditory tasks than the visual tasks (*M* = 12.00, *SD* = 14.741 and *M* = 7.97, *SD* = 12.017, respectively), *t*(35) = 2.615, *p* < .05 (see Figure 21), *d* = .30.

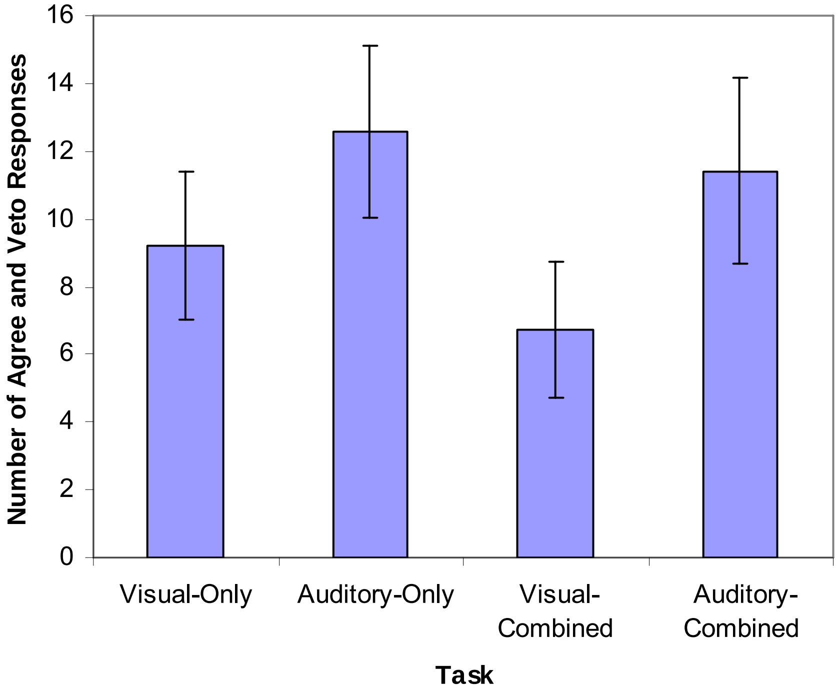


Figure 21. Mean number of times participants clicked on 'Agree' or 'Veto' during each task.

Although there was no significant main effect for AC, *F*(1, 34) = 1.177, *p* > .192, there appeared to be a trend for participants with high AC (*M* = 12.806) to click on 'Agree' or 'Veto' more often than those with low AC (*M* = 7.264), *d* = .53 (see Figure

23). The duty cycles, or patterns of LOA changes during the course of each task, can be viewed for low and high participants in the visual-only task in the appendices of this document (Appendices O and P, respectively), auditory-only task (Appendices Q and R, respectively), visual-combined task (Appendices S and T, respectively), and auditory-combined task (Appendices U and V, respectively).

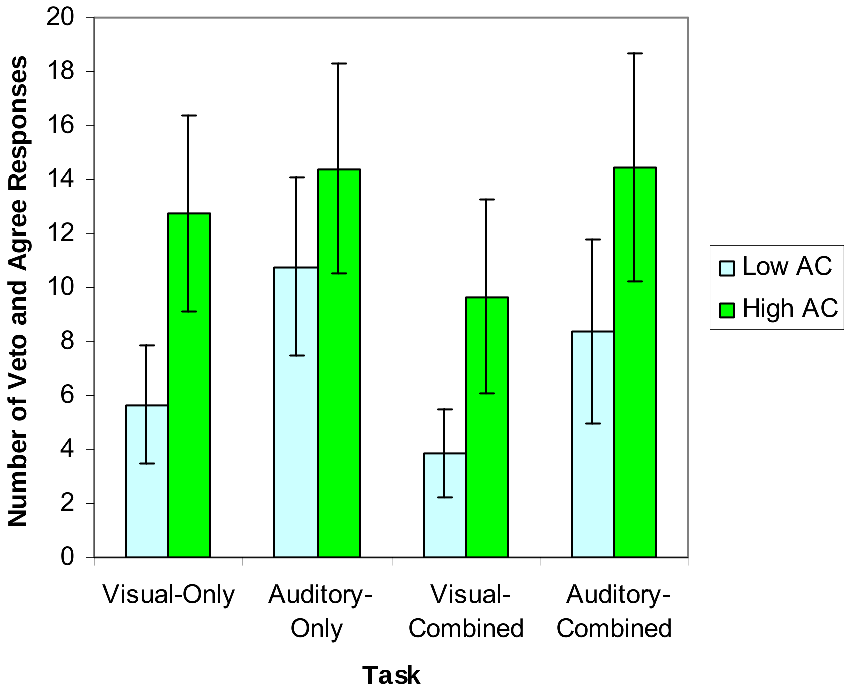


Figure 22. Mean number of times high AC and low AC participants clicked on 'Agree' or 'Veto' during each task.

During the single task condition, the operators were told that other task (the auditory or the visual) was fully automated, but that if they suspected a machine malfunction they should use the Reset Button to recalibrate the machine. This button was established to make sure that operators were still aware of the other task. There was no actual malfunctions in the automation, which was 100% reliable in the fully machine condition. Seven button presses were registered for the Reset Button: during the visual-only task, one participant with low AC pressed it once, while one participant with high AC pressed it twice. During the auditory-only task, one operator with low AC pressed it once, while another operator with low AC pressed it three times.

Signal Detection Performance - d' (Sensitivity)

A hit refers to the case in which the stimulus was a target and the participant correctly indicated that it was a target. A miss is when the stimulus was a target, but the participant erroneously indicated that it was a non. A false alarm is when the stimulus was a non, but the participant erroneously indicated that it was a target. A correct rejection is when the stimulus was a non, and the participant correctly indicated that it was a non. The hit rate is thus the proportion of targets correctly identified, while the false alarm rate is the proportion of nons erroneously identified as targets.

d' is a measure of sensitivity, or the participant's ability to correctly identify the stimuli as a targets or nons. d' integrates both the hit and false alarm rates, and is one of the most commonly used measured of sensitivity in detection theory (Green & Swets, 1966). Perfect sensitivity is characterized by a hit rate of 1 and a false alarm rate of 0. High sensitivity is indicated by a d' which is greater than 0. Low sensitivity is characterized by a hit rate and false alarm rate that approximate equivalency and is indicated by a d' that approximates 0. d' thus typically ranges from 0 to 4.65, which is considered to be an effective ceiling value (Macmillian & Creelman, 2005).

A 2 (Modality) x 2 (Load) x 2 (without-with Machine Aid) within-subjects repeated measured ANOVA was conducted. AC was the between-subjects variable. The ANOVA table (see Table x) summarizes the main effects and interactions involving d'.

Table 6. Analysis of Variance for d' (Sensitivity)

Source	*df*	*F*	η^2	*P*
		Between subjects		
Attentional Control (AC)	1	.666	.019	.420
S within-group error	34	(1.128)		
		Within subjects		
Modality	1	52.343**	.606	.000
Modality x AC	1	.071	.002	.792
Modality x Load	1	1.891	.053	.178
Modality x Load x AC	1	5.622*	.142	.024
Modality x S within group error	34	(1.435)		
Load	1	3.611	.096	.066
Load x AC	1	3.637	.097	.065
Load x S within group error	34	(.449)		
Machine use	1	4.651*	.120	.038
Machine x AC	1	.950	.027	.337
Machine x Modality	1	1.122	.032	.297
Machine x Load	1	1.917	.053	.175
Machine x Load x Modality	1	2.007	.056	.166
Machine x Modality x AC	1	.494	.014	.487
Machine x Load x AC	1	.319	.009	.576
Machine x Load x Modality x AC	1	1.109	.032	.300
Machine x S within group error	34	(.385)		

Note. Values enclosed in parentheses represent mean square errors. *S* = subjects.

**p* < .05. ***p* < .0001.

There was a main effect for modality, such that the d' in the auditory tasks (*M* = 2.163, *SD* = .572) was significantly higher than the d' for the visual tasks (*M* = 1.143, *SD* = .552), *F*(1, 34) = 52.343, *p* < .0001, *d* = 1.82 (see Figure 24). Post-hoc comparisons between tasks were then made. Tukey's HSD ($q = 4.563$, $t = 3.227$) revealed that there was a significant interaction in the operator-alone d' between the visual-only and auditory-only tasks $t(34) = 6.099$, $p < .05$, such that the operator's d' of the auditory-only task was higher than the operator's d' of the visual-only task ($M = 2.036$, $SD = .991$, and $M = .902$, $SD = .779$, $d = 1.35$). In the dual-task condition, the operator-alone d' of the auditory-combined task was significantly higher than that of the visual-combined task ($M = 2.193$, $SD = .898$, and $M = 1.169$, $SD = .735$, respectively), $t(34) = 5.554$, $p < .05$, $d = 1.25$.

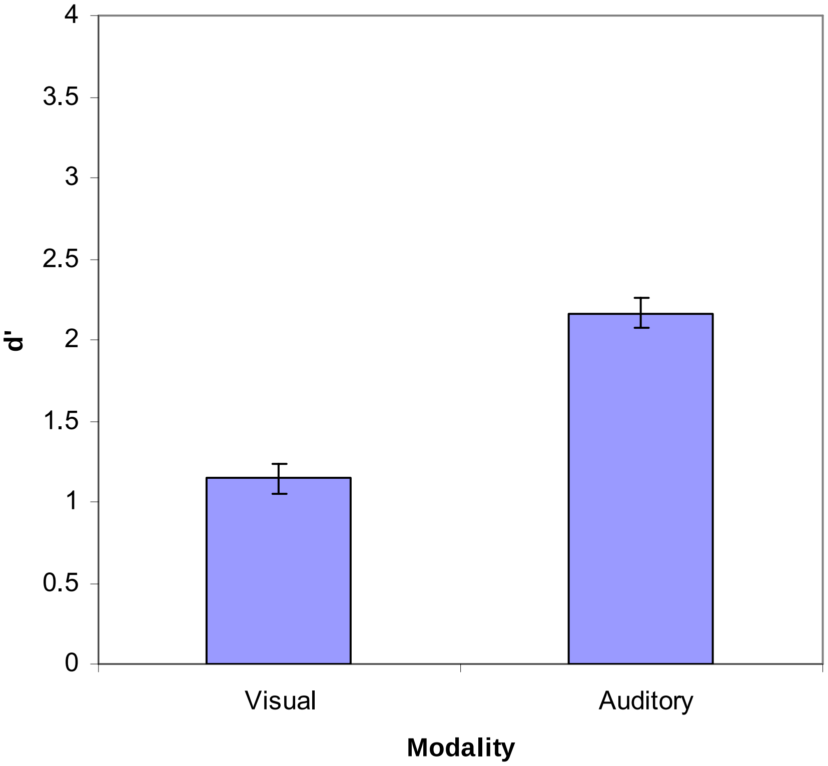


Figure 23. Mean d' as a function of modality.

There was a main effect for use of the machine's decision-aid, as it significantly improved d' across tasks and modality, $F(1, 34) = 4.651$, $p < .05$. Across the visual modality, there was a significant effect for use of the machine decision aid, $t(35) = 2.141$, $p < .05$, such that the operator's d' was significantly when the machine was used ($M = 1.251$, $SD = .589$) than the operator's d' without the machine aid ($M = 1.036$, $SD = .659$), $d = .34$. Across the auditory modality, however, there was no significant improvement for using the machine aid ($p > .05$). There was no significant improvement in d' in using the machine decision aid in the visual-only task, $t(35) = .241$, $p > .05$.

There was not significant effect main effect for load, $F(1, 34) = 3.611$, $p > .05$. However, there was a trend for the d' in the combined task to be higher than the d' in the single-task conditions ($M = 1.729$, $SD = .750$, and $M = 1.579$, $SD = .66$, respectively, $d = .21$).

There was a significant interaction among attentional control, modality, and load, $F(1, 34) = 5.622$, $p < .05$ (see Figure 25). Post hoc analyses using Tukey's HSD ($q = 4.563$, $t = 3.227$) were then conducted to compare the tasks within each AC group.

For participants with high AC, the overall d' in the auditory-combined task was significantly higher than that of the auditory-alone task ($M = 2.311$, $SD = .764$, and $M = 1.931$, $SD = .449$, respectively), $t(17) = 2.183$, $p < .05$, $d = .71$. The overall d' in the auditory-only task was also significantly higher than the d' in the visual-only task ($M = 1.931$, $SD = .764$, and $M = .951$, $SD = .614$, respectively), $t(17) = 4.119$, $p = .001$, $d = 1.81$. The overall d' in the auditory-combined task was also significantly higher than the visual-combined task d' ($M = 2.311$, $SD = .449$, and $M = 1.173$, $SD = .618$, respectively), $t(17) = 5.897$, $p < .0001$, $d = .358$.

For participants with low AC, The overall d' in the auditory-only task was significantly higher than that of the overall visual-only task (*M* = 2.357, *SD* = .766, and *M* = 1.077, *SD* = .453, respectively), *t*(17) = 6.959, *p* < .0001, *d* = 2.36. The overall d' in the auditory-combined task was also significantly higher than that of the visual-combined task (*M* = 2.060, *SD* = 1.059, and *M* = 1.373, *SD* = .756, respectively), *t*(17) = 2.406, *p* < .05, *d* = .92. This did not exceed Tukey's critical t of 3.227, however, the Cohen's *d* was .747.

There was no significant interaction between AC, modality, load, and machine use, *F*(1, 34) = 1.109, *p* > .05. There was a trend for the d' of low AC participants to be higher than that of high AC participants in conditions of machine use in most tasks (see Figure 24): operator-alone d' in the visual-only (*M* = .949, *SD* = .577, and *M* = .855, *SD* = .955, respectively, *d* = .23), machine-aided d' in the visual-only task (*M* = 1.205, *SD* = .460, and *M* = 1.047, *SD* = .688, respectively, *d* = .49), operator-alone d' in the auditory-only task (*M* = 2.311, *SD* = .742, and *M* = 1.762, *SD* = .942, respectively, *d* = 1.05), machine-aided d' in the auditory-only task (*M* = 2.403, *SD* = .867, and *M* = 2.101, *SD* = .736, respectively, *d* = .49), operator-alone d' in the visual-combined task (*M* = 1.318, *SD* = .802, and *M* = 1.021, *SD* = .651, respectively, *d* = .52), and machine-aided d' in the visual-combined task (*M* = 1.427, *SD* = .728, and *M* = 1.325, *SD* = .793, respectively, *d* = .20). However, in the auditory-combined task, high AC participants had a higher mean d' than those with low AC in both the operator-alone condition (*M* = 2.270, *SD* = .487, and *M* = 2.115, *SD* = 1.188, respectively, *d* = .45) and the machine-aided condition (*M* = 2.351, *SD* = .548, and *M* = 2.005, *SD* = .990, respectively, *d* = .89).

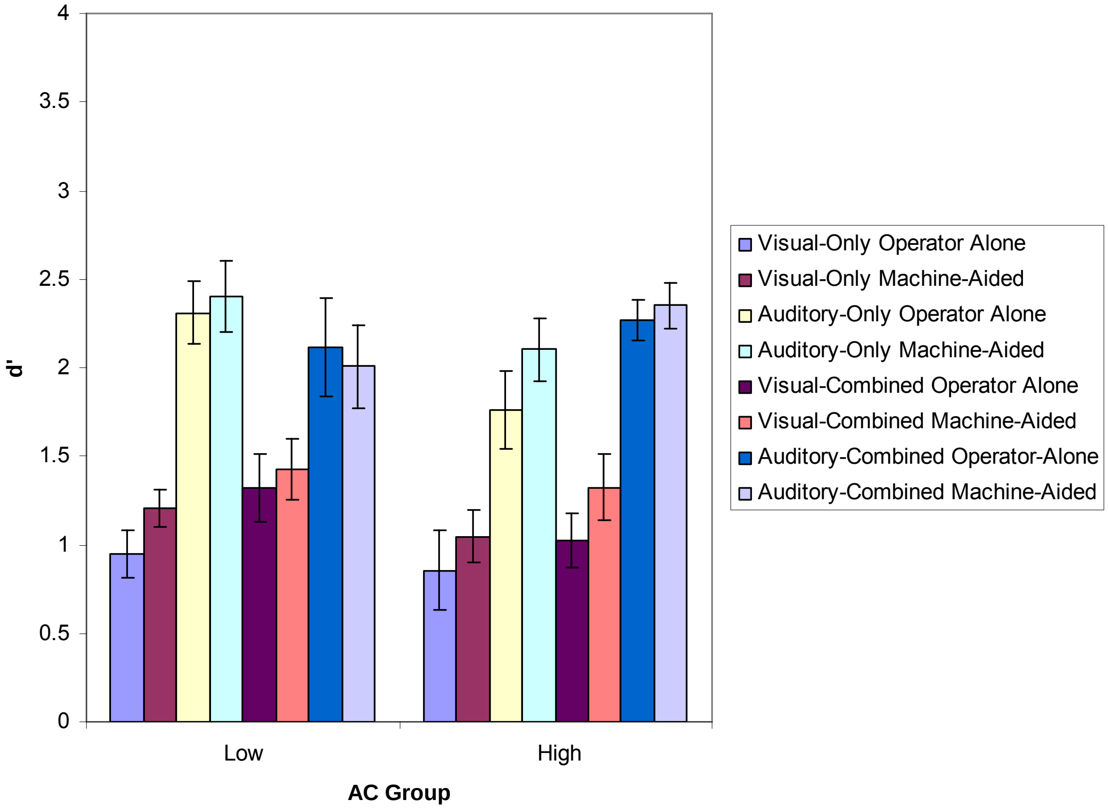


Figure 24. Operator-alone and machine-aided d' in each task grouped by AC.

Response Bias - β (Beta)

β refers to the operator's response bias in identifying stimuli as targets or nons. Macmillian and Creelman (2005), refer to it as a particular favoritism towards a response type. A high β indicates that the observer is more likely to categorize stimuli as 'nons' and have more misses; this suggests a more cautious bias, while low β is the tendency to classify stimuli as 'targets', and commit more false alarms, suggesting more impulsive response behavior (Matthews et al., 2000).

The value of β ranges from 0 to greater than 1. Values in which β > 1 indicate that the individual tends to classify more stimuli as nons (the miss rate exceeds the false alarm rate), resulting in more misses than hits. Conversely, values in which β < 1 indicate that the individual tends to classify more stimuli as targets (the false alarm rate exceeds the miss rate), thereby resulting in more false alarms than hits. When β = 1, there is no response bias, and targets and nons are classified as such according to the actual ratio of targets to nons in the environment (the false alarm and miss rates are equal).

A 2 (Modality) x 2 (Load) x 2 (Machine Aid) within-subjects repeated measures ANOVA was conducted. AC was a between-subjects variable. The ANOVA table (see Table 7) summarizes the main effects and interactions involving β.

Table 7. Analysis of Variance for β (Beta)

Source	*df*	*F*	η^2	*p*
		Between subjects		
Attentional Control (AC)	1	8.185**	.194	.007
S within-group error	34	(1.175)		
		Within subjects		
Modality	1	8.448**	.199	.006
Modality x AC	1	6.408*	.159	.016
Modality x Load	1	.000	.000	.989
Modality x Load x AC	1	.136	.004	.714
Modality x S within group error	34	(.788)		
Load	1	1.612	.045	.213
Load x AC	1	1.392		.246
Load x S within group error	34	(.696)		
Machine use	1	4.145	.109	.050
Machine x AC	1	5.760*	.145	.022
Machine x Modality	1	4.096	.108	.051
Machine x Load	1	1.826	.051	.186
Machine x Load x Modality	1	4.752*	.123	.036
Machine x Modality x AC	1	3.408	.091	.074
Machine x Load x AC	1	.772	.022	.386
Machine x Load x Modality x AC	1	2.304	.063	.138
Machine x S within group error	34	(.332)		

Note. Values enclosed in parentheses represent mean square errors. *S* = subjects.

**p* < .05. ***p* < .01.

There was a significant main effect for AC, *F*(1, 34) = 8.185, p < .01, such that participants with high AC had a significantly higher mean β than those with low AC (*M* = 1.286, *SD* = .491, and *M* = .920, *SD* = .229, respectively, *d* = .49). More specifically, participants with high AC had a significantly higher mean operator-alone β than those with low AC, *t*(34) = 3.361, p < .005, (*M* = 1.4364, *SD* = .620 and *M* = .908, *SD* = .244, respectively, *d* = 1.20). There was no significant difference in machine-aided β between those with low and high AC (p > .05).

There as a significant interaction between AC and modality, *F*(1, 34) = 6.408, *p* <.05. Tukey's HSD (*q* = 3.820, *t* = 2.701) revealed that in the auditory modality, those with high AC had a significantly higher mean β than those with low AC (*M* = 1.570, *SD* = .749, and *M* = .940, SD = .386, respectively), *t*(34) = 3.169, p < .005, *d* = 1.19. However, there was no such difference within the visual modality *t*(34) = .822, *p* > .05. See Figure 26 for a depiction of the mean β in each modality grouped by AC.

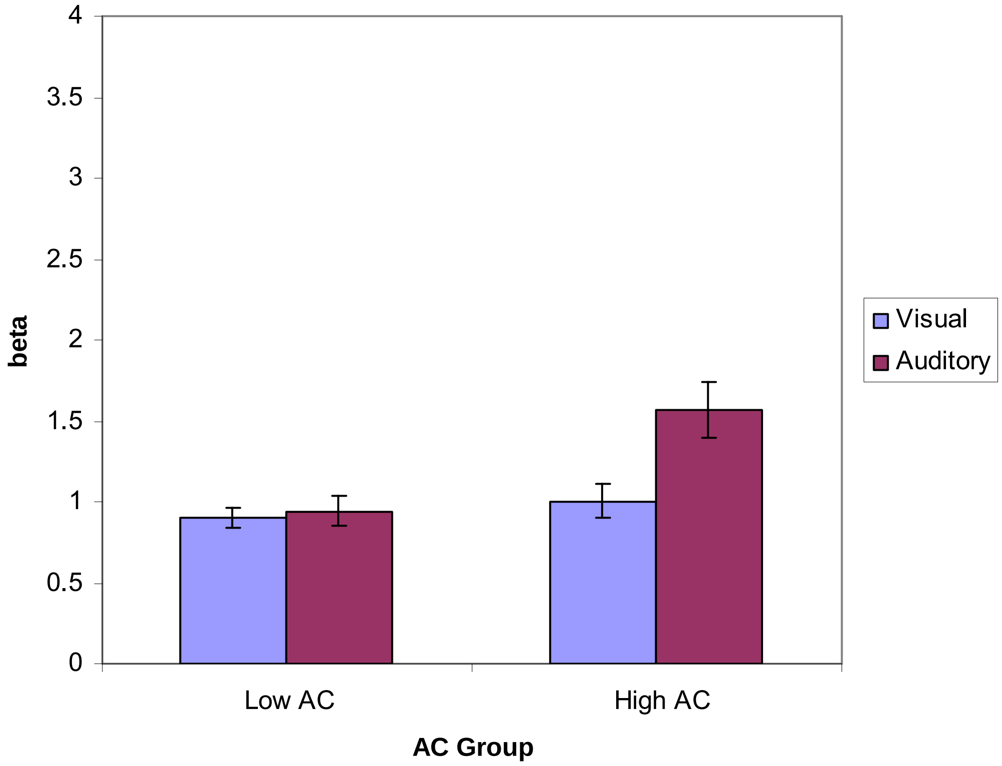


Figure 25. Mean β by AC group and modality.

For high AC participants, there was a significant interaction among modality, machine aid, and load, *F*(1, 17) = 4.754, *p*<.05. Post hoc pairwise comparisons using Tukey's HSD (*q* = 4.858, *t* = 3.435) indicated that in the auditory-combined task, the operator-alone β was marginally higher than that of the machine-aided β (*M* = 2.142, *SD* = 1.693, and *M* = 1.287, *SD* = .607, respectively), *t*(17) = 2.193, p < .05, however, the critical t did not exceed Tukey's critical t. The Cohen's *d* was .71.

The differences among the operator-alone and machine-aided βs for the visual-only, auditory-only, and visual portion of the combined tasks were not significantly different (p<.05). See Figure 27 for a depiction of the mean βs for each task and condition of machine use for the low and high AC groups.

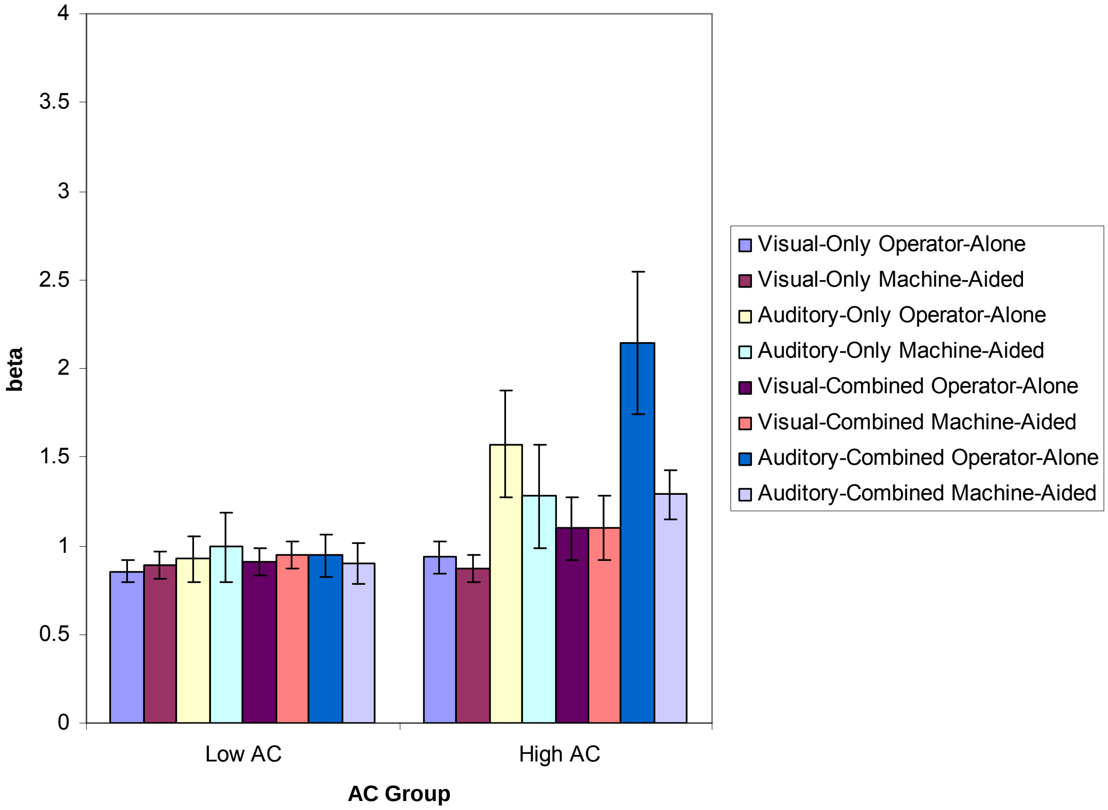


Figure 26. Mean β for by task for both operator-alone and machine-aid, grouped by AC.

There was a main effect for modality in that the auditory β (*M* = 1.255, *SD* = .669) was significantly higher than the visual β (*M* = .951, *SD* = .365), *F*(1,34) = 8.448, *p* < .01.

Use of the machine decision-aid also marginally reduced β across tasks, *F*(1, 34) = 4.145, *p* = .50. The mean β of the operator-alone was 1.172 (*SD* = .536), while the mean machine-aided β was 1.034 (*SD* = .400), *d* = .29. A post hoc analysis using Tukey's HSD (*q* = 3.820, *t* = 2.701) revealed a marginal difference between operator-alone and machine-aided β in the auditory modality, *t*(35) = 1.958, *p* = .058, such that the machine-aided β was lower (*M* = 1.115, *SD* = .606) than the operator-alone β (*M* = 1.396, *SD* = .948), *d* = .35.

There was no main effect for load, *F*(1, 34) = 1.612, *p* > .05.

There was a significant interaction among load, modality, and use of the machine aid, *F*(1, 34) = 4.752, *p* < .05. Post-hoc comparisons were then made using Tukey's HSD (*q* = 4.563, *t* = 3.27). In the auditory-combined task, the operator-alone β was significantly higher than the machine-aided β (*M* = 1.542, *SD* = 1.274, and *M* = 1.094, *SD* = .581, respectively), *t*(35) = 2.184, *p* < .05. Although not significant, Cohen's *d* for this comparison was .461. Figure 28 depicts this interaction in the single-task condition (a), and in the combined condition (b). Figure 29 depicts the mean operator-alone and machine-aided β information for each task.

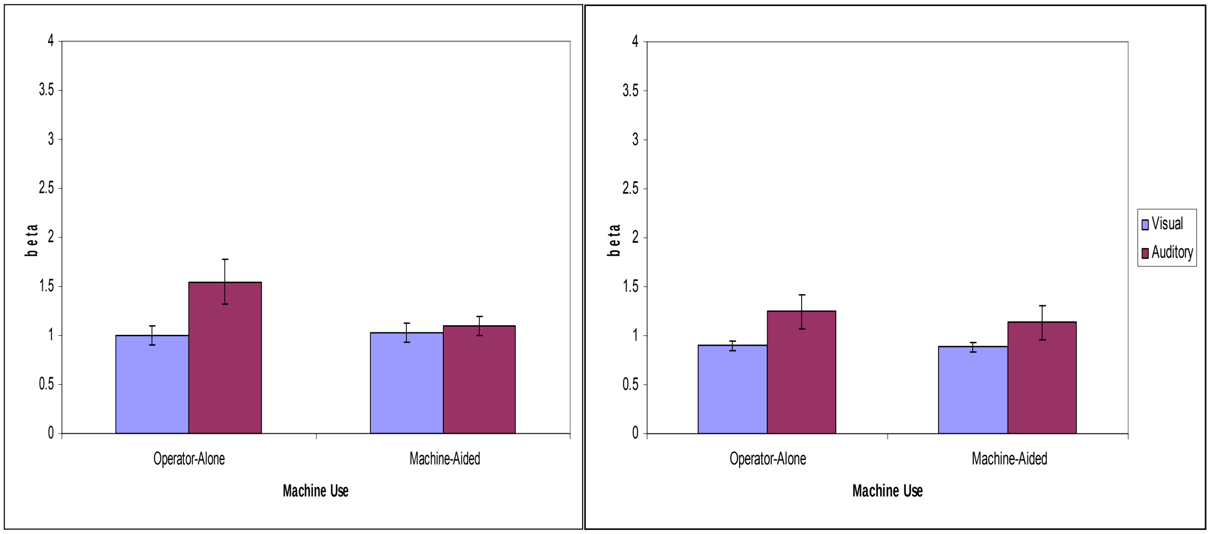


Figure 27. Interaction among machine use, modality, and load. Graph depicts the interaction in the single-task condition (a). Interaction among machine use and modality load. Graph depicts the interaction in the combined condition (b).

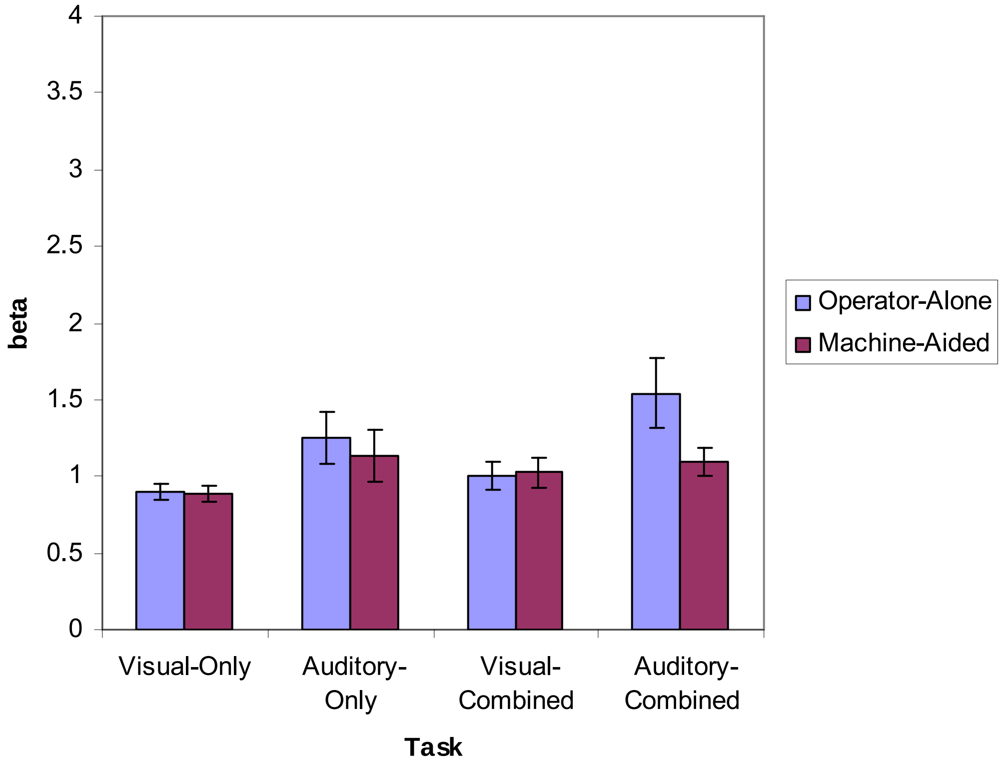


Figure 28. Mean operator-alone and machine-aided beta in each task.

Subjective Workload Ratings

Workload measurements were taken three times over the course of the experiment: after the visual only condition, after the auditory condition and after the combined condition. A 3-way (Task) within-subjects repeated measures ANOVA was used. AC was a between subjects variable. The main effects and interactions for the workload ratings can be seen in Table 8.

Table 8. Analysis of Variance for Workload Ratings

Source	*df*	*F*	η^2	*p*
		Between subjects		
Attentional Control (AC)	1	2.193	.061	.148
S within-group error	34	(1128.460)		
		Within subjects		
Task	2	9.908*	.226	.000
Task x AC	2	.707	.020	.497
Task x S within group error	68	(109.298)		

Note. Values enclosed in parentheses represent mean square errors. *S* = subjects.

**p* < .0001.

A priori tests of simple effects found that the high AC group reported significantly higher overall workload in the visual-only task than the low AC group (*M* = 51.37, *SD* = 11.948, and *M* = 41.95, *SD* = 12.909, respectively), *t*(34) = 2.271, *p* < .05, *d* = 1.11. Although not significant, there was a trend for the high AC group to report higher overall workload in the auditory-only task (*M* = 46.97, *SD* = 7.710, and *M* = 44.07, *SD* = 11.597, respectively, *d* = 1.08) and in the combined task (*M* = 53.94, *SD* =

48.19, *SD* = 13.858, and *M* = 48.19, *SD* = 13.180, respectively, *d* = .59). See Figure 30 for a graph representing the overall workload reports in each task by AC group.

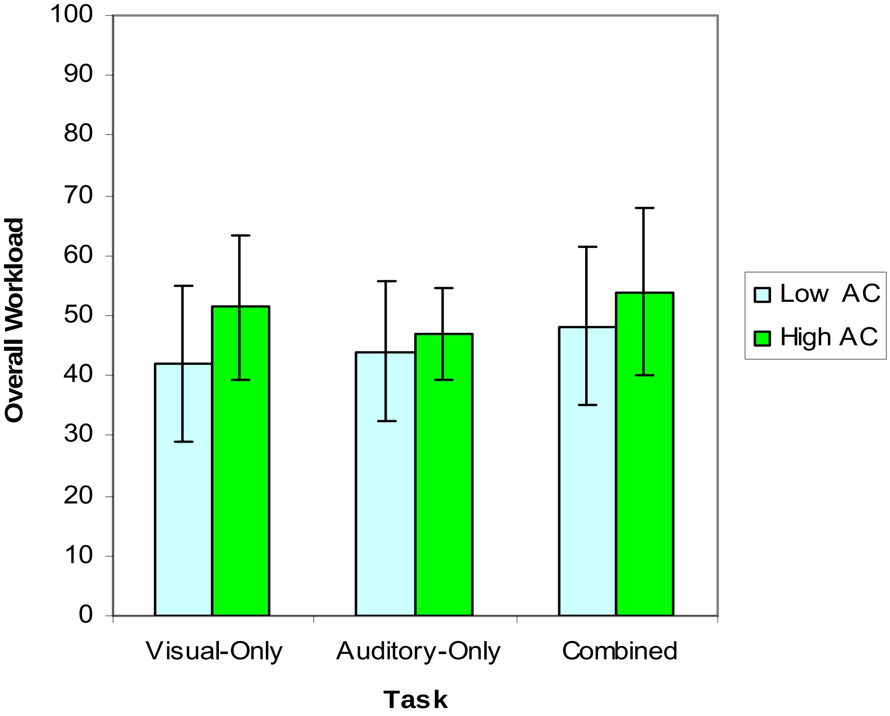


Figure 29. Mean overall workload rating in each task for low and high AC groups.

There was no significant main effect for AC across subscales or any significant interactions involving AC across subscales (all *p* > .05). However, a priori comparisons found that those with high AC reported significantly higher effort in the visual-only task than those with low AC (*M* = 73.22, *SD* = 19.907, and *M* = 54.56, *SD* = 21.864, respectively), *t*(34) = 2.678, *p* < .05, *d* = 1.33. Those with high AC also experienced significantly higher overall workload in the visual-only task than those with low AC (*M* = 51.37, *SD* = 11.948, and *M* = 41.95, *SD* = 12.909, respectively), *t*(34) = 2.271, *p* < .05, *d* = 1.11. All other differences were insignificant (*p* > .05). The difference between effort in the combined task was marginally higher for those with high AC (*M* = 73.17, *SD* = 20.048) than those with low AC (*M* = 61.56), *SD* = 17.813), *p* = .057, *d* = .82. See Figure 31 for a graph depicting the pattern of workload scores for those with low and high AC.

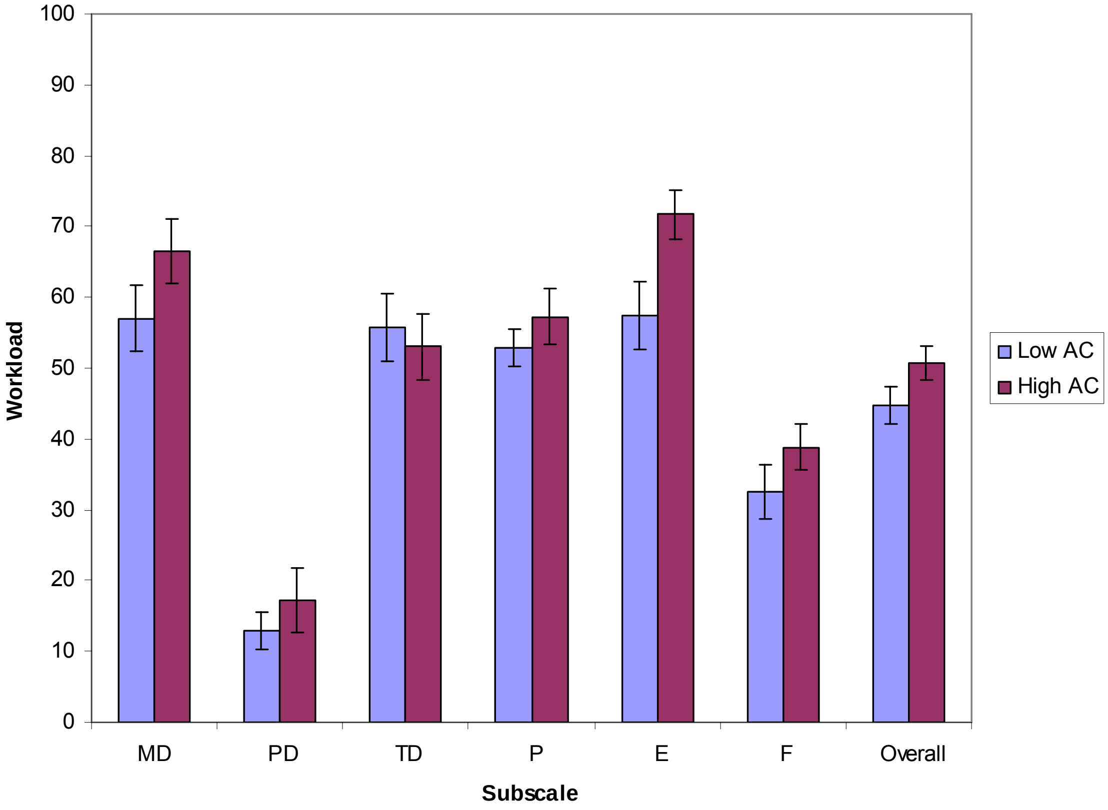


Figure 30. Pattern of workload scores across subscales for low and high AC groups. *Note.* MD = Mental Demand, PD = Physical Demand, TD = Temporal Demand, P = Performance, E = Effort, F = Frustration, and O = Overall

There was a significant main effect for task, *F*(2, 68) = 7.116, $p < .005$ (see Figure 32). The combination task had significantly higher overall workload ratings than the auditory-only task ($M = 51.06$, $SD = 13.643$ and $M = 45.52$, $SD = 9.816$, respectively), $t(1, 35) = 3.340$, $p < .005$, $d = .47$. The combination task also had significantly higher overall workload rating than the visual-only task ($M = 46.66$, $SD = 13.643$), $t(35) = 3.186$, $p < .005$, $d = .32$. There was no significant difference between the overall mean workload in the visual-only and auditory-only tasks ($p > .05$).

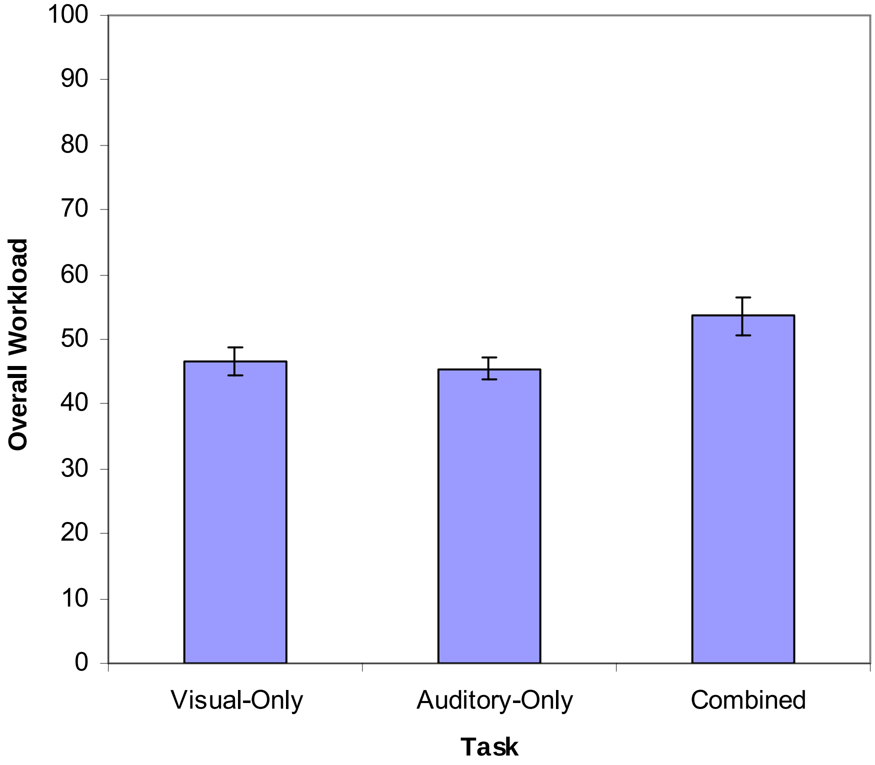


Figure 31. Overall workload in each task.

There was a significant interaction between task and scale, *F*(12, 408) = 7.464. $p < .0001$. Tukey's HSD ($q = 5.098$, $t = 3.461$) was performed on all post hoc analyses comparing the each of the visual-only and auditory-only subscales. See Table 9 for a summary of the significant differences between the subscales. For performance, the auditory-only performance scale was significantly higher than that of the visual-only scale (*M* = 64.28, *SD* = 15.031 and *M* = 47.36, *SD* = 17.978, respectively), *t*(1, 35) = 5.593, *p* < .0001, *d* = 1.02. The mental demand in the combination task was significantly higher than that of the auditory-only task, (*M* = 67.50, *SD* = 20.891, *M* = 56.58, *SD* = 20.656, respectively), *t*(1, 35) = 4.671, *p* < .0001, *d* = .53. The physical demand in the combination task was significantly higher than that of the auditory-only task, (*M* = 19.28, *SD* = 20.453, and *M* =11.42, *SD* = 12.100, respectively), *t*(1, 35) = 3.948, p < .0001, *d* = .47. Performance in the auditory-only task was significantly higher than that of the combined task, (*M* = 64.28, *SD* = 15.031, and *M* = 53.61, *SD* = 17.919, respectively), *t*(1, 35) = 3.963, *p* < .0001, *d* = .65. Frustration in the combination task was significantly higher than that of the auditory-only task, (*M* = 39.89, *SD* = 18.338, and *M* = 28.75, *SD* = 14.916, respectively), *t*(1, 35) = 3.648, p < .005, *d* = .67.

Table 9. Significant Differences Between Subscales by Task.

	Visual-Only	Auditory-Only	Combined
Mental Demand	*M* = 61.11 *SD* = 22.817	*M* = 56.58 *SD* = 29.656	*M* = 67.50 *SD* = 20.89
		....	....
Physical Demand	*M* = 14.44 *SD* = 21.121	*M* = 11.42 *SD* = 12.100	*M* = 19.28 *SD* = 20.453
		....	....
Temporal Demand	*M* = 54.67 *SD* = 22.726	*M* = 49.67 *SD* = 24.194	*M* = 58.75 *SD* = 23.249
Performance	*M* = 47.36 *SD* = 17.978	*M* = 64.28 *SD* = 15.031	*M* = 53.61 *SD* = 17.919
	....	....	....
Effort	*M* = 63.89 *SD* = 22.678	*M* = 62.44 *SD* = 20.854	*M* = 67.36 *SD* = 19.596
Frustration	*M* = 38.50 *SD* = 21.788	*M* = 28.75 *SD* = 14.916	*M* = 29.89 *SD* = 18.338
		- - - - - - -	- - - - - - -
Overall	*M* = 46.66 *SD* = 13.156	*M* = 45.52 *SD* = 9.816	*M* = 51.06 *SD* = 13.643
		- - - - - - -	- - - - - - -

Note. Significantly different pairs are indicated with lines. Each line patterns indicates a specific *p* value as follows:

— — — — -	$p < .01$
- - - - - - -	$p < .005$
....	$p < .0001$

Discussion

The main purpose of Experiment 2 was to determine the role of attentional behavior in preference for various levels of task automation. Participants with low and high attentional control (AC: Derryberry & Reed, 2002) were recruited. Each participant assumed the role of the operator of an automated machine decision-aid, and completed a visual-only, auditory-only, and audiovisual combined task in order to assess modality effects as well as single and dual-task effects. These perceptual tasks involved making temporal discriminations, in which participants identified auditory and visual stimuli presented for the shorter duration of time as ‘targets’ and those of the longer duration as ‘nons’. In each single-task condition, the operator was responsible for ensuring either only the lines were correctly identified in the visual-only task, or the tones were correctly identified in the auditory-only task. In the dual-task condition, the operator was responsible for ensuring that both auditory and visual stimuli were correctly identified. The dual-task condition was thus designed to be more demanding, as four information sources were to be monitored: the left and right sides of the screen for the visual task, and the left and right headphones for the auditory task. Task load was manipulated to yield variability in subjective workload reports. Single-task conditions required the participant to make the discriminations for stimuli of one modality while stimuli of the other modality were to be fully processed by the machine and thus ignored by the operator.

The operator could complete each task using the LOA of choice, of which there were four gradations ranging from full operator control to mostly machine control with the operator serving as the supervisor of the machine’s decisions in identifying the stimuli. The machine decision-aid was 90% accurate in its identifications in order to encourage operator trust, but preclude full reliance. Because AC is characterized by differences in engaging in, focusing upon, and disengaging upon relevant information, it was hypothesized that individuals with low AC would prefer greater LOAs to aid them, notably in dual-task conditions when all four information sources were to be monitored and both modalities taxed. Those with high AC, on the other hand, would be more adept at managing stimuli from both modalities in dual-task conditions and would therefore need less automation to aid them.

It was also suspected that the high AC group would outperform the low AC group; however, the low AC group may be able to improve their performance by choosing higher LOAs. The machine-aid available to them was 90% reliable, and thus presents itself as a potentially effective decision-aid. Thus, operators with different attentional capacities would be able to optimize their performance by selecting the LOA which best complimented their needs as recognized.

Signal detection theory (SDT: Green & Swets, 1966) was employed to assess the participants' sensitivity (d') to target and non stimuli, as well as response bias (β) towards identifying these stimuli.

Workload measures were also expected to be impacted by AC. Particularly, the high AC group was hypothesized to report lower reports of subjective workload due to their inherent ability to process information from multiple modalities and filter information when necessary. Additionally, the dual-task condition was hypothesized to induce higher subjective workload reports.

Preferences for LOAs

Statistically, LOAs 1, 2, and 4 were equally preferred across tasks and levels of AC. Thus, participants from both AC groups found them all to be equally usable and useful. Participants equally preferred various amounts of assistance from the machine, ranging from no assistance (i.e., LOA 1), to supervising the machine as it made all of the stimuli discriminations and intervening only when a machine decision error was suspected (i.e., LOA 4). LOA 2 was also frequently chosen; participants may have preferred completing the tasks at this level as they received a type of performance feedback, even if 90% accurate. Commentary from participants of both the low and high AC groups corroborated this notion; they reported that the tasks were difficult and generally trusted their own decisions, and receiving suggestions from a machine which was correct 90% reassured them of their choices. They also expressed appreciation for being able to change their response based on the machine's suggestion. This indicates a general attitude that participants were confident with their own abilities enough to accept the primary responsibility for the task, but also preferred having a second opinion which was typically correct.

Participants of both the low and high AC groups spent minimal time at LOA 3, Equal Sharing, when given the choice of four different LOAs. LOA 3 functionality was based upon laterality, in which two different LOAs were in operation on each side: all stimuli on the left side of the screen and left earphone were conducted using LOA 2, in which the operator had the primary identification responsibility, while all stimuli on the right side of the screen and right earphone were conducted using LOA 4, in which the machine had the primary identification responsibility. This places LOA 3 in the middle of the range of the human-machine responsibility scale; LOA 3 required even less operator responsibility than LOAs 1 and 2, both of which were more preferred than LOA 3. It was also less popular than LOA 4, in which minimal operator responsibility was required as a supervisory role was assumed. This suggests that LOA 3 may have been least preferred for reasons other than workload and responsibility. One possible reason may be that operators did not wish to filter stimuli, such that they were required to attend to certain stimuli (i.e., left fields) while ignore others of equal perceptual intensity and modality type (i.e., right fields). That is, certain lines and tones had to be selectively ignored, while other equally salient lines and tones had to be selectively processed. The human perceptual mechanisms may not be naturally inclined towards or trained for lateral filtering; few common daily tasks require processing information only on the left side of the perceptual field while ignoring those on the right. LOA 3, which incorporated processing and responding schemes from two different LOAs may have also placed an additional information processing and strategy-development demand upon the operator, where as the other LOAs used the same strategy for all stimuli, regardless of laterality. That is, operators had to learn and use two different strategies in LOA 3, one for the left side and one for the right side, whereas the other LOAs only required one strategy for both sides. As mentioned in the literature review, planning strategies for task management and attentional allocation can increase workload (Tulga & Sheridan, 1980); LOA 3 may have thus doubled the strategy planning demand. Olson and Sarter (2000) state that automation can also increase workload if the operator is excessively taxed by the requirement to interact with and supervise it; LOA 3 may have been such a case, and operators readily recognized it as such.

Modality effects for LOA preference were unanticipated; however, Experiment 2 revealed that there was a trend for more time to be spent in LOA 2 during the auditory tasks than during the visual tasks. This indicates that participants willingly assumed more task control in the auditory modality than in the visual modality. Furthermore, the higher d' and lower workload ratings for the auditory tasks suggests that participants may have had a more positive overall experience completing the auditory tasks than the visual tasks and were thus more inclined to be engaged in the tasks.

Conversely, LOA 4 was more preferred in the visual tasks than in the auditory tasks; thus, the visual tasks were thus more likely to be offloaded to the machine's responsibility. Additionally, the visual tasks were characterized by lower d' and higher workload ratings compared to the auditory tasks. These findings suggest that participants may have favored completing the auditory tasks more than the visual tasks, and they thus assigned more of the visual task responsibility to the machine.

There was a trend for LOA 4 to be used more often in the dual-task conditions than in the single task conditions, and more specifically, there was also a trend to spend more time in LOA 4 in the visual-combined task than the visual-only task. This suggests that participants were more inclined to use automation to offload the more demanding load of the dual-task condition. This is corroborated by higher reports of overall workload in the dual-task condition than the single-task conditions. Additionally, there was a trend for the d' in the combined task to be higher than the d' in the single-task conditions. This suggests that the decision-making functionality associated with the higher LOAs was used to compensate for the higher workload and difficulty associated with dual-tasking. The use of automation in this experiment thus fulfilled one of its primary uses, which is to serve as a decision-making agent to relieve high workload in a multitasking environment (Madni, 1988).

No significant main AC effects were found in terms of LOA preference, however, there were trends for those with low AC to spend more time at LOA 4 than those with high AC, collapsed across tasks. Also, those with high AC spent more time at LOA 1 than did those with low AC. Thus, lower AC levels were conducive to the preference for more automation and higher AC levels were conducive to inclinations towards less automation. This supports the prediction that AC is inversely proportional to automation use. However, the preference for LOA 2 was roughly equivalent for both low and high AC groups; indicating that both types of individuals liked using LOA 2, a moderately-low amount automation.

There was also a trend for low AC participants to use LOA 4 in the visual tasks, which were found to be more difficult post-hoc, than in auditory tasks, which were associated with a higher d’ (see Figure 33 for a comparison). This indicates that the low AC participants recognized that the visual tasks were more challenging and thus preferred to have more machine assistance to offload some of the task demand.

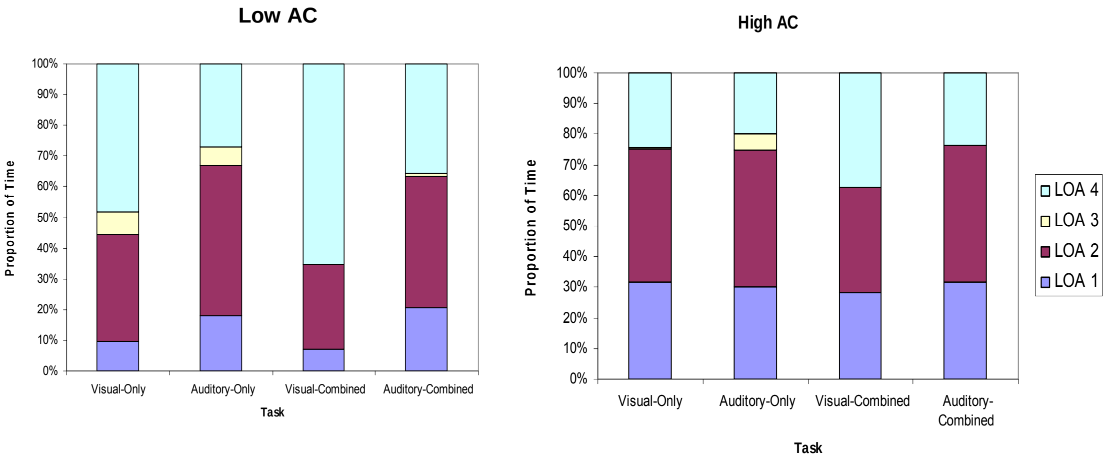


Figure 32. Overall proportion of time spent at each LOA for each task for the low and high AC groups.

There was a trend for those with low AC to change LOAs more frequently than those with high AC in all four task types. This may be because those with high AC were able to focus upon the algorithms of a task and learn how to use it more efficiently without being distracted, and thus could more quickly settle upon a specific LOA. Low AC participants may have had more difficulty settling upon an LOA, or perhaps grew bored with one and were inclined to try another. They may have also focused less upon the nature of the stimuli, making less notice of task demands, and focused more upon the various control modes available to them.

Another consideration is the finding from Experiment 1 that high AC is directly proportional to desirability of control at a correlation of .35. Thus, those with low AC would have lower desirability of control, and may therefore be more inclined towards

LOA 4. Those with high AC, however, may have more desirability of control and thus prefer LOA 1.

Signal Detection Performance - d' (Sensitivity)

Although not hypothesized, participants were generally more sensitive to auditory stimuli than visual stimuli, as evidenced by the higher d' for the auditory tasks. Despite prior psychophysical equation of these two discrimination types by Szalma et al. (2004), the auditory discriminations were easier than the visual discriminations. This d' discrepancy also perpetuated despite the fact that the temporal difference between target and non tones was less than that of the target and non lines (i.e., 50 ms difference in the auditory modality vs. 125 ms in the visual modality, respectively). This modality effect in d' also replicated the results by Szalma et al. 2004). Additionally, this modality d' difference persisted in individuals with both low and high AC.

The machine decision-aid also improved the operators' d' in the visual modality, indicating that use of the automation was helpful and effective as intended. However, it did not increase d' in the auditory modality, likely due to the fact that the auditory d' was already significantly higher than that of the visual modality and perhaps approached a ceiling effect in which additional aid was no longer useful to the operator. Participants may have already felt comfortable making the auditory discriminations; this is in light of the finding that participants generally used lower LOAs in the auditory modality than in the visual modality. Thus, there may have been less need for automation in the auditory tasks. The visual tasks, which were more challenging, influenced participants to implore the machine for more assistance, as is to be expected in light of the common use of automation to relieve operator demand for tasks that are difficult in terms of sensory thresholds (Wickens, 1992; Lee and See, 2004).

There was not a significant main effect for load; the single- and dual-task conditions had comparable mean d' levels. A few participants commented that they preferred the audiovisual combined task because attending to all stimuli types kept them more task-engaged, which they enjoyed. This may be because it was easier to process all salient incoming information rather than actively filter out a major source of stimulation (e.g., the selectively ignoring the tones during the visual-only task). The requirement of actively filtering out a source of salient information also involves developing a strategy for allocating resources. According to Tulga & Sheridan (1980), it is often the inertia effect that prevails: even if the operator acknowledges that performance could improve by changing the current strategy, the extra effort required to generate and adopt a new strategy is often thought to be 'excessive', and consequently the path of least resistance prevails. There may also be other individual differences in divided and selective attention which have yet to be assessed. Matthews et al. (2000) suggest that there is no general ability to divided attention or timeshare performance of concurrent tasks. Instead, it may be more task-specific; the current tasks may not have tapped such differences. Future investigations can examine this relationship by presenting various levels of task demand which may relate to attentional control, such as integrating and comparing information from a greater number of sources, elevating the event rate, or perhaps taxing only one modality by presenting two or more concurrent visual tasks rather than a task from each modality.

There was also a trend for low AC participants to have a higher mean d’ than high AC participants for most of the tasks: the visual-only, auditory-only, and visual-combined. This trend persisted in both operator-alone and machine-aided d’; this may be partially attributed for by the tendency for low AC participants to spend more time at higher LOAs, wherein they received more machine assistance than those with high AC, who tended to use LOA 1 more frequently. This may indicate that those with low AC compensated for their attentional failures by using more automation; consequently, they were able to raise their performance levels to those of the high AC group. The high AC group outperformed the low AC group in the auditory-combined task, however. In observing the usage pattern of LOAs by task for each AC group, it can be seen that low AC participants used LOA 4 more in the visual-combined task, and thus their d’ for this task was comparable to those with high AC. However, the low AC group used LOA 4 less in the auditory-combined task, and here their performance was lower than those with high AC. Had the low AC group continued to use LOA 4 more often in the auditory-combined task, their performance may have been equal to or marginally higher than that of the high AC group. The low AC participants may have therefore needed more machine assistance in this task, but did not recognize it as such.

Response Bias - β (Beta)

Auditory tasks had higher mean β than visual tasks, and furthermore, use of the machine decision-aid also marginally reduced the operators’ β across tasks. The interaction among load, modality, and use of the machine decision aid was driven by a reduced mean β as a result of using the machine, and was only seen in the combined condition. However, these effects on β were likely governed by differences between AC groups. The low AC group had consistently low-to-neutral β across task variables, but the high AC group showed greater variability in mean β (see Figure 26).

The high AC group was found to have a significantly higher overall mean β than the low AC group across all tasks. This was specifically found to be true when parsing machine use effects; the operator-alone β was higher for the high AC group than for the low AC group, however, this effect disappeared when comparing the machine-aided β for the low and high AC groups. Thus, use of the machine aid significantly reduced the β for only the high AC group; the low AC group was not as influenced by the machine aid in terms of β. This may be a result of the high AC participants' ability to attend to integrate more sources of information simultaneously; they may have been more adept at responding to the stimuli as well as the machine's input.

The difference in β for AC groups was likely driven by modality effects. The mean β of the high AC group was only higher than that of the low AC group in the auditory modality. High AC participants were thus more likely to report identifying nons in the auditory tasks than they were in the visual task. There were no such differences in modality for the low AC group. Again, use of the machine aid moderated this effect. Machine use attenuated this β difference in the high AC group; the machine-aided β was lower than the operator's initial β in the auditory tasks. Thus, while the high AC participants were more likely to report more targets, the machine aid was effective in reducing their β to approximate neutrality. The high AC group was more affected by the machine aid; their more effective attentional behavior may have enabled them to integrate the machine input, which was an additional source of stimulation in the task and could therefore demand additional attentional resources.

The low AC group, however, consistently reported more targets than the high AC group. The low AC group was also more consistently neutral across modalities and also across machine use conditions. This may have been the result of the fact that the low AC group had more of an inclination towards using LOA 4, in which the machine answer was the default. The machine's β was always 1, identifying an equal number of targets and nons. Thus, by using a higher proportion of machine answers, low AC participants were more likely to have a mean β closer to 1.

Note, however, that the mean β of the low AC group was consistently neutral and sometimes less than 1, and in most of the tasks, the operator-alone β was lower than the machine-aided β. Thus, low AC participants were generally neutral, and occasionally erred on the side of identifying more stimuli as targets, thereby committing more hits, but also more false alarms. Matthews et al. (2000) describe this response pattern as being more impulsive: indicating more affirmative responses when negative responses are correct. It is thus of interest to seek links between low AC and impulsivity and bias towards perceiving targets. Impulsivity is a component of extraversion (Eysenck, 1967). The link between AC and extraversion, especially in regards to response bias impulsivity, thus presents itself as an avenue for future investigation.

AC has been shown to correlate with trait anxiety (TA). Derryberry & Reed (2001) report a correlation of -.55, and Experiment 1 of the current study established a correlation of -.45. Those with high AC thus have lower TA, and conversely, low AC individuals have higher TA. Anxiety effects are often task-specific (Eysenck, 1981), such as difficult tasks, short-term and working memory tasks, and secondary tasks in dual-tasking (Darke, 1988; Eysenck, 1992, 1997).

With respect to the Cognitive Failures Questionnaire (CFQ: Broadbent, Cooper, FitzGerald, & Parkes, 1982), anxious individuals tend to rate themselves as prone to errors, possibly because of a tendency towards self-focused attention can lead to cognitive failures (Matthews & Wells, 1988). Anxious individuals are thus distracted by their own thoughts (Carver, Peterson, Follansbee, & Scheier, 1983). Aware of their own attentional failure tendency, the low AC group may have been somewhat more concerned with missing targets. The high AC group may have hesitated and thought more before identifying the stimuli as targets, and tended to be more conservative.

High TA individuals also have a bias towards processing the threatening aspects of an event (French, 1992), such as in responding to threatening words, bias in dichotic listening (Matthews & MacLeod, 1985), and bias in visual search (MacLeod & Matthews, 1986). Thus, in Experiment 2, if the term ‘target’ did in fact suggest a threat connotation, then those with low AC, who are more trait anxious, may then tended to focus on the threat of targets and have a reduced criterion and consequently lower β for identifying targets compared to the high AC group.

According to Eysneck’s hypervigilance theory (1992), anxious individuals first scan the environment for a threat, then focus on channels where threat stimuli are detected. This can cause distraction from the task at hand and then consequent performance failure, unless the task requirement involves processing threat stimuli. This may partially explain why there was no observed performance (d’) failure in the current task among those with low AC (and thus high TA). In fact, those with low AC often outperformed those with high AC. In accordance with the hypervigilance theory, low TA participants efficiently focused upon the threat sources, which were the screen and headphones, and were therefore likely to detect targets. This is of course, presuming that the term ‘target’ was perceived as threatening. Further examination to identify this relationship may involve defining the target as a threat in order to assess the potentially differential reaction between low and high AC and TA groups. Because the instruction set of the current experiment did not suggest a threat connotation with the term ‘target’, neutral response biases were anticipated.

Subjective Workload Ratings

The abbreviated version of the NASA-TLX (Hart & Staveland, 1988) was administered after each of the three tasks to assess patterns in subjective mental workload as influenced by task load and AC level. As hypothesized, the combined task yielded higher workload ratings than both of the single task conditions. This indicates that the higher task load was more subjectively demanding than the lower loads. The combined task had significantly higher subjective workload ratings than the single-tasks in terms of most of the specific subscales: mental demand, physical demand, and frustration.

The temporal demand was also statistically equivalent at the mid-range in all three tasks, indicating that the event rate was challenging yet feasible.

Participants indicated that they felt they performed the best on the auditory-only task, followed by the combined task, and finally, they indicated that they performed the worst in the visual-only task. This is congruent with the overall lower mean d’ associated with the visual tasks. Conversely, and also congruent with its associated mean d’, the auditory-only task received consistently low workload ratings.

Contrary to the hypotheses, those with high AC had a trend of reporting higher workload than those with low AC on almost all of the subscales; it was anticipated that these individuals would have less difficulty with all three tasks due to their ability to attend to various stimuli types, especially in the combined task. Significantly higher reports among the high AC participants occurred in the overall workload rating in the visual-only task, and effort in the visual-only task. The reasons for this may include the tendency for those with high AC to spend more time in LOA 1 than those with low AC; LOA 1 places all identification responsibility upon the operator. Conversely, the low AC participants spent more time in LOA 4 overall. Another possible explanation for this is that the high AC participants may have been attending to the nature of the task more so than the low AC participants. Thus, the high AC group may have been more engaged in the tasks, and therefore more prone to noticing their own feelings of workload.

Implications for Results

In Experiment 2, the degree of automation was flexible as per the preferences of the operators. The machine decision-aid was effective, as participants switched LOAs according to preference and use of the machine aid improved d’ and rendered a more neutral response bias than the operator’s initial inclination. While three of the available LOAs were equally preferred, and task performance was variable and neither subjected to floor nor ceiling effects, operators may not have necessarily accurately judged their need for various LOAs. In this experiment, the aim was determining preference rather than need for automation. However, a discrepancy between need and preference has been noted by Morrison and Rouse (1986), as the operator’s ability to assess the need for automation may not always be realized. It is therefore not determinable whether the LOAs chosen by the operators in Experiment 2 truly optimized task performance. Sensitivity to the task stimuli may have been a function of LOA choice, however, operators may have reacted differently in the case of fixed LOAs. Additionally, response bias and subjective workload may have been variable as a function of AC and modality because of the LOA flexibility. The potential effect of LOA flexibility versus rigidity was thus deemed an important variable in operator-machine interaction. Experiment 3, which is detailed in Chapter 6, was then conducted using a similar methodology, with the one exception that participants were assigned to a either a low or high LOA for task completion.

CHAPTER 6: EXPERIMENT 3 – LOW AND HIGH ATTENTIONAL CONTROL IN FIXED LEVELS OF AUTOMATION

A third experiment was conducted to observe performance, response bias, and workload effects in response to fixing the automation at a certain level rather than providing the operators with a choice. Half of the participants were assigned to a low LOA, while others were assigned to a high LOA. Additionally, both low and high AC groups were assigned to each of these LOAs to investigate the influence of attentional behavior.

Hypotheses

Based on the original hypotheses generated by the literature review as well as the results of Experiment 2, the following a priori predictions were made.

1. The d' of the higher load (combined) will exceed that of the lower load (single-task), possibly because the higher load encourages task engagement.
2. As observed in Experiment 2, there will be an interaction between load and modality, such that the d' of the auditory-combined task will be higher than that of the auditory-only task, and the d' of the visual-combined task will be higher than that of the visual-only task.
3. The higher LOA 4 will result in a higher machine-aided d' than the lower LOA 2.
4. In LOA 2, the machine-aided d' will be higher than operator-alone d'.
5. Tasks drawing from the auditory modality will have a higher d' than the visual modality.
6. Use of the machine aid will further raise the d' in each modality.

7. In LOA 2, the machine-aided β will be closer to 1 than operator-alone responses. Thus, the machine-aided β will be lower than the operator-alone d’.
8. There will be a main effect for AC, such that the β will be higher for high AC participants than for low AC participants.
9. There will also be a significant interaction between AC and modality, such that in the auditory modality, those with high AC will have a significantly higher mean β than those with low AC.
10. Those with high AC will also have higher β for the auditory modality than for the visual modality.
11. Subjective workload reports will be lower for those using LOA 4 than 2.
12. High AC participants will also report higher workload than those with low AC.
13. The visual-only task will have higher workload ratings than the auditory alone task, and the combined task will result in higher workload ratings than the single-task conditions.
14. Participants will agree with or veto the machine more often in LOA 2 than LOA 4.
15. Participants will agree with or veto the machine more often in the auditory modality than visual modality.
16. Participants will agree or veto with the machine more in LOA 2 than in LOA 4. This will also vary as a function of modality, in that they will also respond to the machine more often in the auditory modality than the visual modality.

Methodology

Experimental Participants

Thirty-two participants were drawn from the pool generated in Experiment 1. Participants with low and high AC scores (44 or less and 56 and greater, respectively) were selected. Recruitment criteria included consent to be contacted for recruitment to subsequent studies (see Appendix F) and Attentional Control scores in the lower and upper quartile ranges as described in the Experiment 1 methodology (see Figure 34). Sixteen participants with low AC were selected, and sixteen participants with high AC were selected. Ages ranged from 18 to 34 (*M* = 21.19, *SD* = 2.693). Low AC scores ranged from 34 to 43 (*M* = 39.686, *SD* = 2.651) and high AC scores ranged from 56 to 67 (*M* = 61.313, *SD* = 2.600). Participants from the low and high AC groups were equally assigned to one of two fixed LOAs: either 2 (mostly manual) or 4 (mostly machine).

The low AC group fixed in LOA 2 consisted of 8 participants: three males and five females. Among the low AC male participants, the mean age was 21 years (*SD* = 1) and the mean AC score was 39.33 (*SD* = 4.619). Among the low AC female participants, the mean age was 22 years (*SD* = 1.225) and the mean AC score was 39.8 (*SD* = 2.588).

The low AC group fixed in LOA 4 consisted of 8 participants: two males and three females. Among the low AC male participants, the mean age was 21.5 years (*SD* = .707) and the mean AC score was 41.5 (*SD* = .707). Among the low AC female participants, the mean age was 19.83 years (*SD* = 1.169) and the mean AC score was 39.17 (*SD* = 2.317).

The high AC group fixed in LOA 2 consisted of 8 participants: three males and five females. Among the high AC male participants, the mean age was 21 years (*SD* = 2) and the mean AC score was 62 (*SD* = 3). Among the high AC female participants, the mean age was 20.6 years (*SD* = .894) and the mean AC score was 62.6 (*SD* = 2.702).

The high AC group fixed in LOA 4 consisted of 8 participants: three males and five females. Among the high AC male participants, the mean age was 24 years (*SD* = 8.718) and the mean AC score was 61.33 (*SD* = .577). Among the high AC female participants, the mean age was 21 years (*SD* = 1.225) and the mean AC score was 59.6 (*SD* = 2.702).

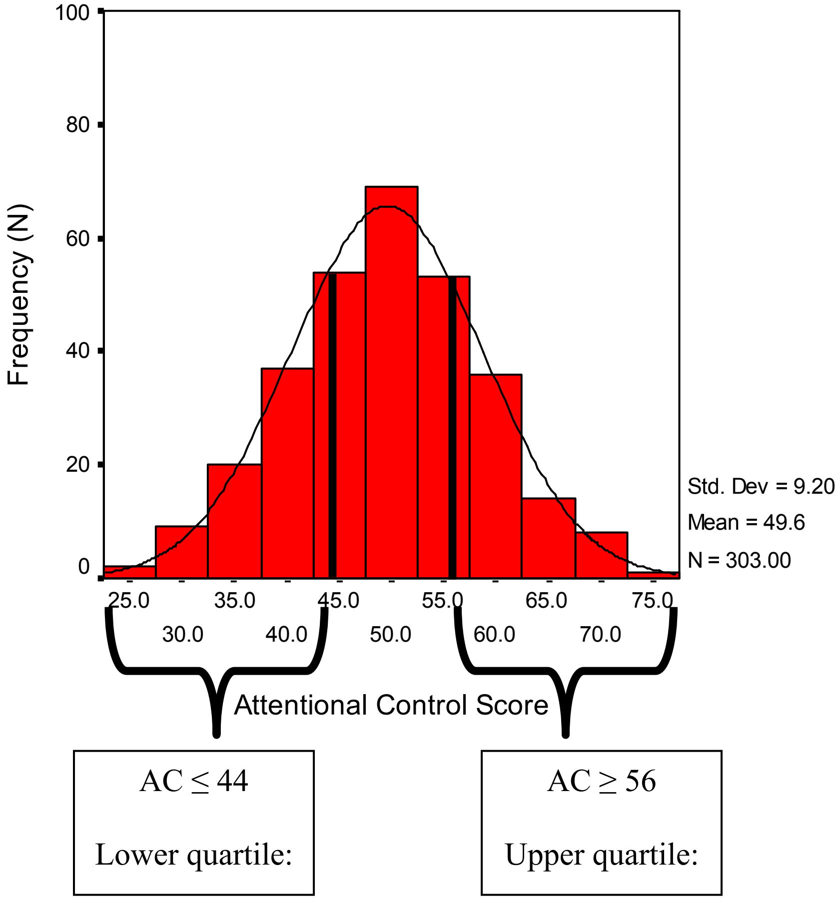


Figure 33. Histogram which features the 'Low AC' and 'High AC' overlay upon the AC distribution. Participants were selected from the high and low AC groups.

Procedure

After a brief description of the experiment by the experimenter, participants completed an informed consent form (appendix L) to affirm that they were at least 18 years of age. The automation program functioned identically to that in Experiment 2, with one exception that participants were not able to switch the LOA during the tasks. Rather, they were randomly assigned to a particular LOA at which to complete the tasks.

Half of the participants (N = 16) were assigned to LOA 2, while the other half of the participants (N = 16) were assigned to LOA 4. Half of the participants assigned to LOA 2 had low AC scores (N = 8, AC ≤ 44 or less), while the other half assigned to LOA 2 had high AC scores (N = 8, AC ≥ 56). Likewise, half of the participants assigned to LOA 4 had low AC scores, while the other half had high AC scores. Thus, there was an equal number of low and high AC participants assigned to each LOA.

As in Experiment 2, they were read instructions from an instruction manual (see Appendix M for instructions for participants fixed in LOA 2 and Appendix N for instructions for participants fixed in LOA 4). They were only trained to use the LOA at which they completed the task. They also practiced the task at the assigned LOA for two minutes before beginning the tasks. They were informed that if they tried to change the LOA in the Visual and Auditory Control Mode menus, nothing would happen and the task would continue at the LOA which had been assigned to them.

As in Experiment 2, each participant completed the same visual-only, auditory-only, and audiovisual combined tasks, using the same stimuli types and response mechanisms for identifying targets and nons. They also completed the abbreviated version of the NASA-TLX after each of the three tasks.

Results

Signal Detection Performance - d' (Sensitivity)

A 2 (Modality) x 2 (Load) repeated measures within-subjects ANOVA was conducted. AC group (2: low and high) and LOA group (2: LOA 2 and LOA 4) were between-subjects variables (see Table 10).

LOA 4 only has a machine-aided component, since the machine answer is the default, as is the nature of this high LOA. Thus there is no operator-alone component in LOA 4. There are no comparisons involving machine use effects for LOA 4. LOA 2 retains both operator-alone and machine aided, since the operator answer is the default and can be altered by machine use. Thus, the following ANOVA compares only the final responses for each LOA group: the machine-aided d’ for LOA 2 and the machine-aided d’ for LOA 4.

Table 10. Analysis of Variance for d' (Sensitivity)

Source	*df*	*F*	η^2	*p*
		Between subjects		
AC group	1	.483	.017	.493
AC x *S* within-group error	28	(.750)		
LOA group	1	4.578*	.141	.041
LOA x *S* within-group error	28	(.750)		
		Within subjects		
Modality	1	144.116**	.837	.000
Modality x AC	1	.350	.012	.559
Modality x LOA	1	.002	.000	.968
Modality x AC x LOA	1	.618	.022	.438
Modality x *S* within-group error	28	(.478)		
Load	1	2.403	.079	.132
Load x AC	1	.035	.001	.853
Load x LOA	1	1.823	.061	.188
Load x AC x LOA	1	1.272	.043	.269
Load x *S* within-group error	28	(.153)		
Modality x Load	1	1.334	.045	.258
Modality x Load x AC	1	1.495	.051	.232
Modality x Load x LOA	1	1.249	.043	.273
Modality x Load x AC x LOA	1	.128	.005	.723

Note. Values enclosed in parentheses represent mean square errors. *S* = subjects.

$*p < .05$. $**p < .0001$.

There was a significant main effect for LOA, *F*(1, 28) = 4.578, $p < .05$, such that the machine-aided d' was higher in LOA 4 (*M* = 2.266, *SD* = .379) than in LOA 2 (*M* = 1.768, *SD* = .688), *d* = .90. See Figure 35 for the difference in mean d' as a function of LOA.

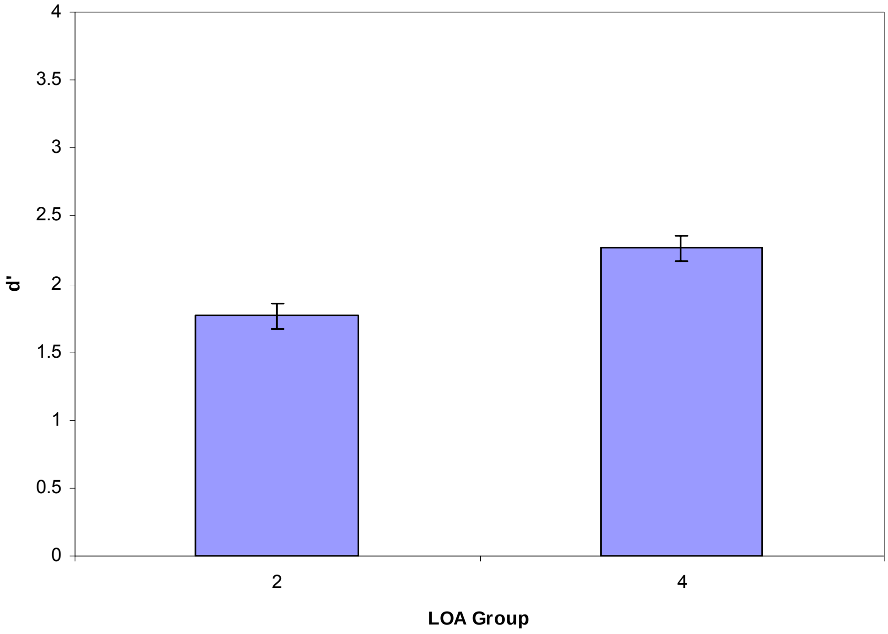


Figure 34. Mean d' as a function of LOA.

There was a significant main effect for modality, *F*(1, 28) = 144.116, $p < .0001$ (see Figure 36). Auditory machine-aided d' (*M* = 2.836, *SD* = .587) was significantly higher than visual machine-aided d' (*M* = 1.369, *SD* = .530), *d* = 2.62.

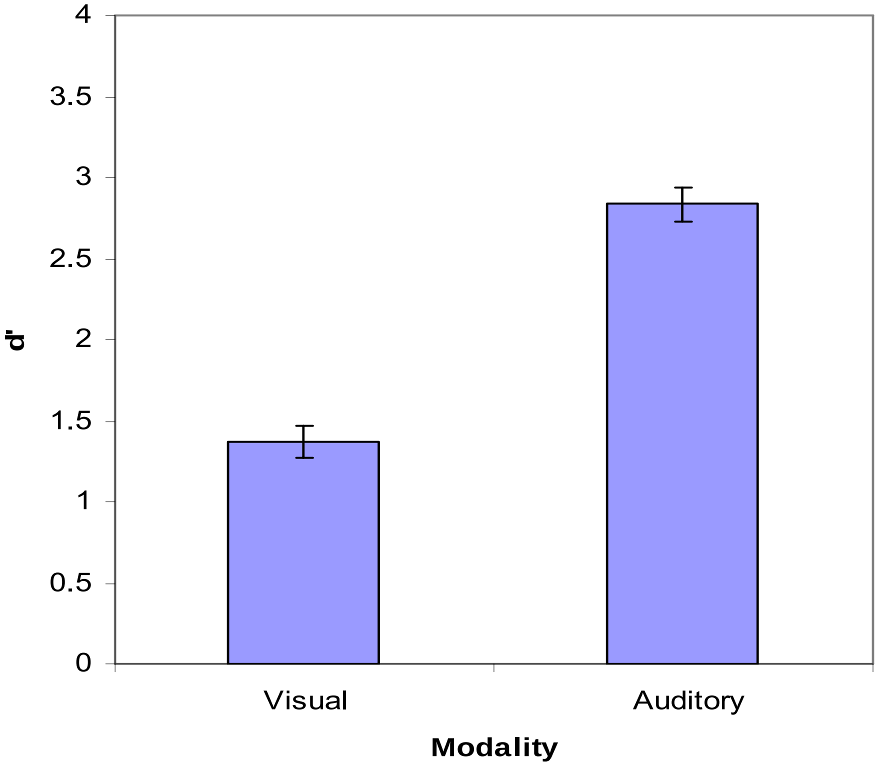


Figure 35. Mean machine-aided d' as a function of modality.

There was no significant main effect for load, $F(1, 28) = 2.403$, $p > .05$.

Machine-aided d' in each task was compared between LOA 2 and LOA 4 using a priori test of simple effects. In the auditory-combined task, the machine-aided d' was higher in LOA 4 than in LOA 2 ($M = 3.185$, $SD = .605$, and $M = 2.683$, $SD = .692$, respectively), $t(30) = 2.182$, $p < .05$, $d = 1.17$. All other pairs were nonsignificant (all $p > .05$).

A separate set of a priori tests of simple effects was done to compare the operator-alone and machine-aided d' in LOA 2 only. Machine-aided d' was significantly higher than operator-alone d', $t(15) = 2.656$, $p < .05$. In LOA 2, the machine-aided d' was significantly higher than the operator-alone d’ in the visual-only task (*M* = 1.197, *SD* = .650, and *M* = .503, *SD* = .598, respectively), *t*(15) = 2.982, *p* < .01, *d* = 1.15. All other pairs were nonsignificant, but there was a trend for machine-aided d’ to be higher than operator-alone d’: the auditory-only task (*M* = 2.666, *SD* = .786, and *M* = 2.388, *SD* = .875, respectively, *d* = .33), the visual-combined task (*M* = 1.208, *SD* = .623, and *M* = .892, *SD* = .660, respectively, *d* = .49), and the auditory-combined task (*M* = 2.683, *SD* = .692, and *M* = 2.606, *SD* = .746, respectively, *d* = .11). See Figure 37 for a graph depicting the mean d’ in LOA 2 by machine use. Also see Figure 38 for mean d’ as a function of LOA group, modality, load, and machine use. This graph shows both operator-alone and machine-aided d’ for both LOAs 2 and 4.

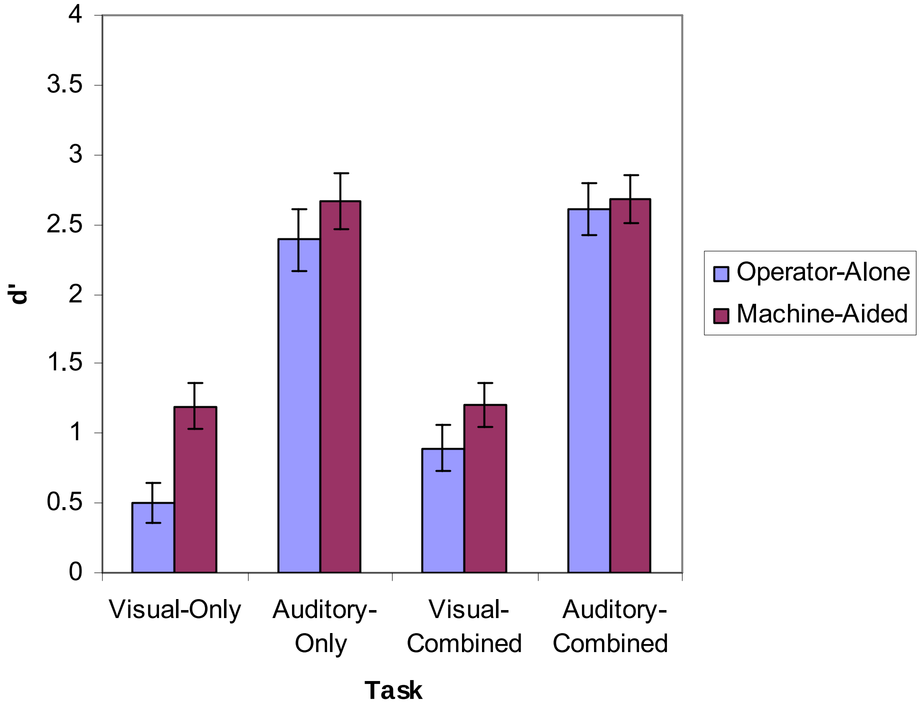


Figure 36. Mean operator-alone and machine aided d’ in each task for LOA 2.

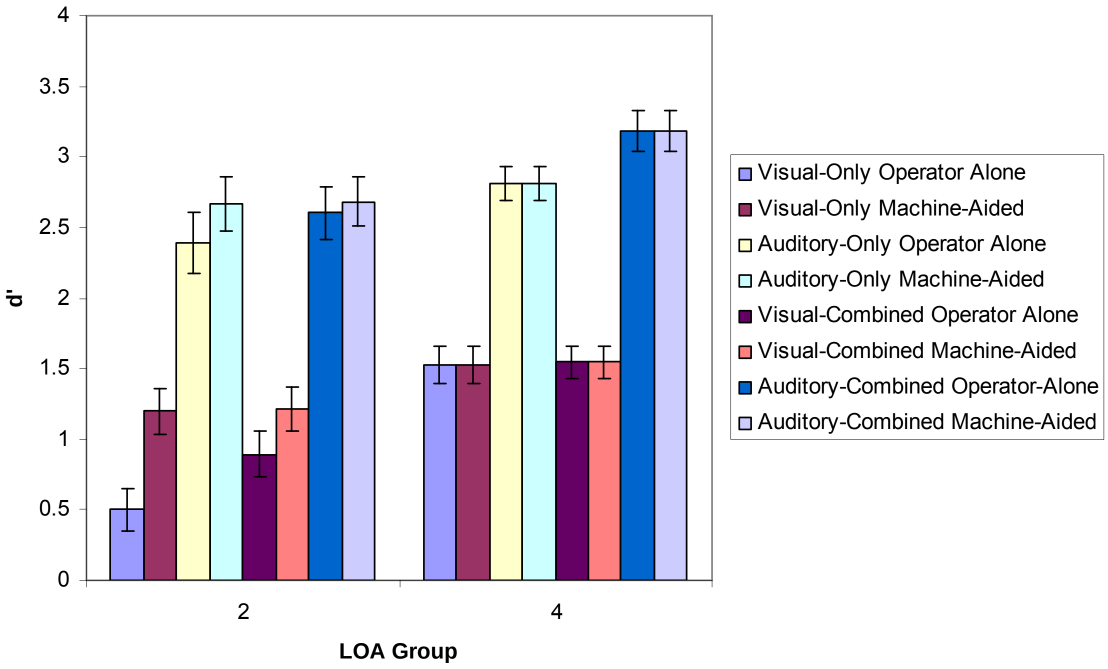


Figure 37. Mean d' as a function of LOA group, modality, load, and machine use.

In terms of machine-aided d', there was no main effect for AC, *F*(1, 28) = .483, *p* > .05. There was a trend for auditory d' to exceed visual d' across both AC and LOA groups (see Figure 39 for the low AC group, and Figure 40 for the high AC group). Tests of simple effects were performed to identify differences in each task between the low and high AC groups in each LOA, however, no significant differences were found.

In the LOA 2 group, the low AC group d' was higher than that of the high AC group in the visual-only task (*M* = 1.446, *SD* = .574, and *M* = .948, *SD* = .658, respectively, *d* = 1.23), in the auditory-only task (*M* = 2.680, *SD* = .646, and *M* = 2.652, *SD* = .952, respectively, *d* = .09), and in the visual-combined task (*M* = 1.299, *SD* = .457, and *M* = 1.117, *SD* = .777, respectively, *d* = .56). The high AC group had a higher d' than the low AC group in the auditory-combined task (*M* = 2.694, *SD* = .603, and *M* = 2.672, *SD* = .814, respectively, *d* = .05).

In the LOA 4 group, the low AC group d' was higher than that of the high AC group in the visual-only task (*M* = 1.558, *SD* = .484, and *M* = 1.485, *SD* = .621, *d* = .21) and in the auditory-combined task (*M* = 3.310, *SD* = .414, and *M* = 3.059, *SD* = .760, respectively, *d* = .86). The high AC group had a higher d' in the auditory-only task (*M* = 2.871, *SD* = .517, and *M* = 2.749, *SD* = .507, respectively, *d* = .33), and in the visual-combined task (*M* = 1.568, *SD* = .379, and *M* = 1.529, *SD* = .560, respectively, *d* = .15).

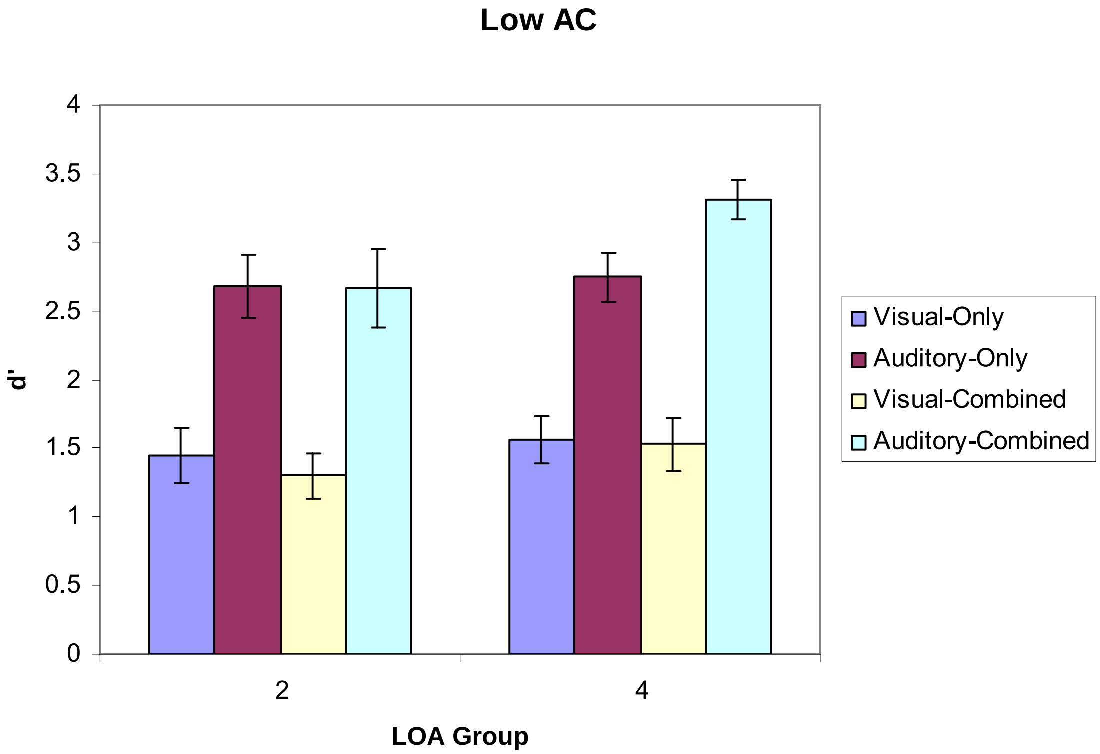


Figure 38. Mean machine-aided d' in each task as a function of LOA Group for the low AC group.

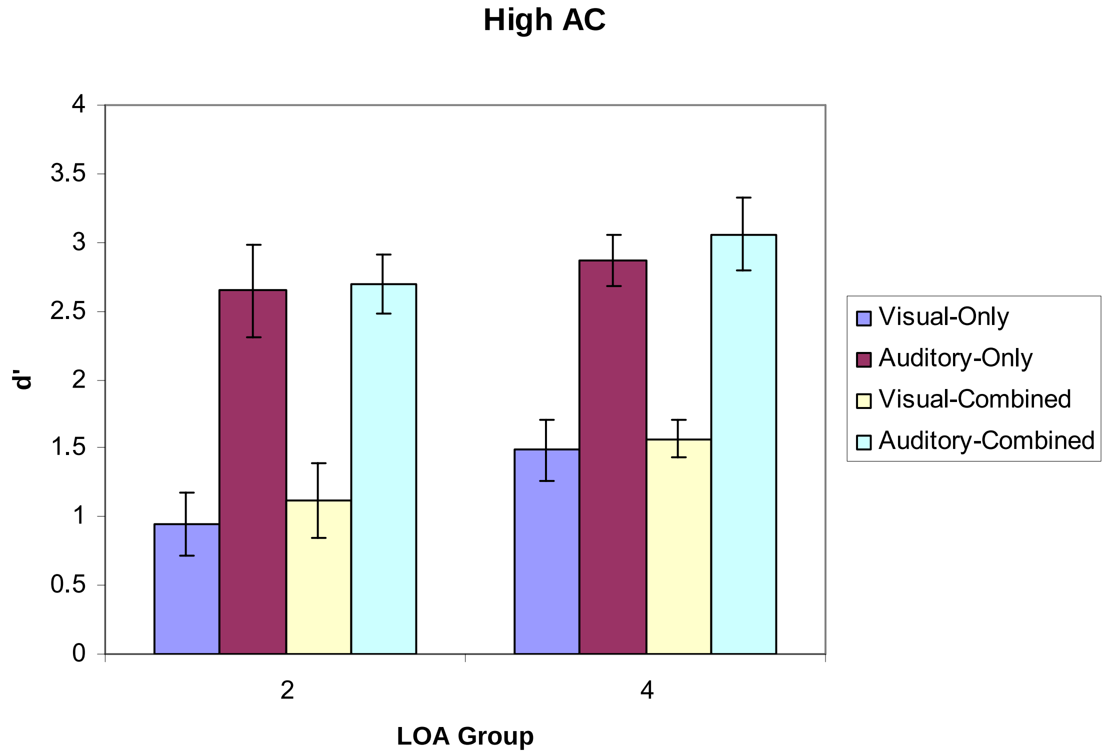


Figure 39. Mean machine-aided d' in each task as a function of LOA Group for the high AC group.

Figure 41 shows the mean operator-alone and machine-aided d' for each task across LOAs.

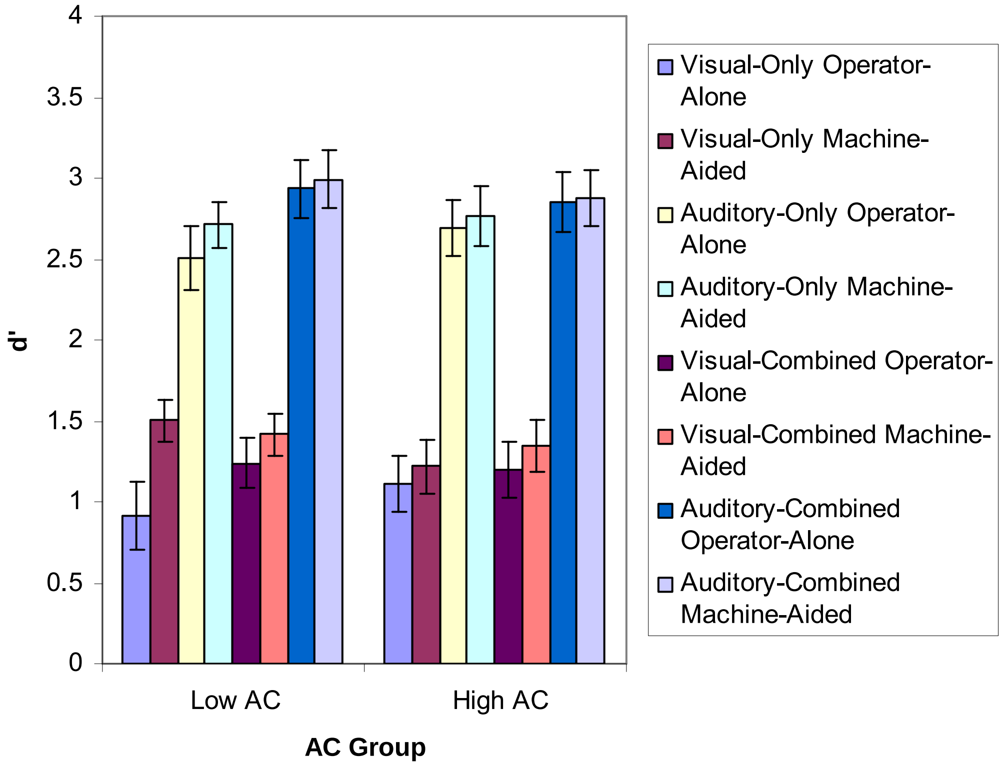


Figure 40. Mean operator-alone and machine-aided d' in each task as a function of AC group.

Response Bias – β (Beta)

A 2 (Modality) x 2 (Load) repeated measures within-subjects ANOVA was conducted. AC group (2: low and high) and LOA group (2: LOA 2 and LOA 4) were between-subjects variables.

As in the d' analyses, the following ANOVA compares only the final response biases for each LOA group: the machine-aided β for LOA 2 and the machine-aided β for LOA 4. Table X features the effects on β resulting from the ANOVA.

Table 11. Analysis of Variance for β (Response Bias)

Source	*df*	*F*	η^2	*P*
		Between subjects		
AC group	1	6.823	.196	.014
AC x *S* within-group error	28	(1.404)		
LOA group	1	3.109	.100	.089
LOA x *S* within-group error	28			
		Within subjects		
Modality	1	.687	.024	.414
Modality x AC	1	10.351	.270	.003
Modality x LOA	1	1.113	.038	.300
Modality x AC x LOA	1	2.969	.069	.096
Modality x *S* within-group error	28	(1.332)		
Load	1	.123	.004	.729
Load x AC	1	.010	.000	.921
Load x LOA	1	.002	.000	.962
Load x AC x LOA	1	.322	.011	.575
Load x *S* within-group error	28	(1.158)		
Modality x Load	1	2.073	.161	.069
Modality x Load x AC	1	1.140	.039	.295
Modality x Load x LOA	1	.009	.000	.925
Modality x Load x AC x LOA	1	(1.015)		

Note. Values enclosed in parentheses represent mean square errors. *S* = subjects.

$^*p < .05$. $^{**}p < .01$. $^{***}\ p < .005$, $^{****}\ p < .0001$

There was a significant main effect for AC, *F*(1, 28) = 6.823, $p < .05$, such that participants with high AC had a higher β than those with low AC (*M* = 1.661, *SD* = .725, and *M* = 1.114, *SD* = .453, respectively, *d* = 1.07). See Figure 42 for the mean machine-aided β by AC group.

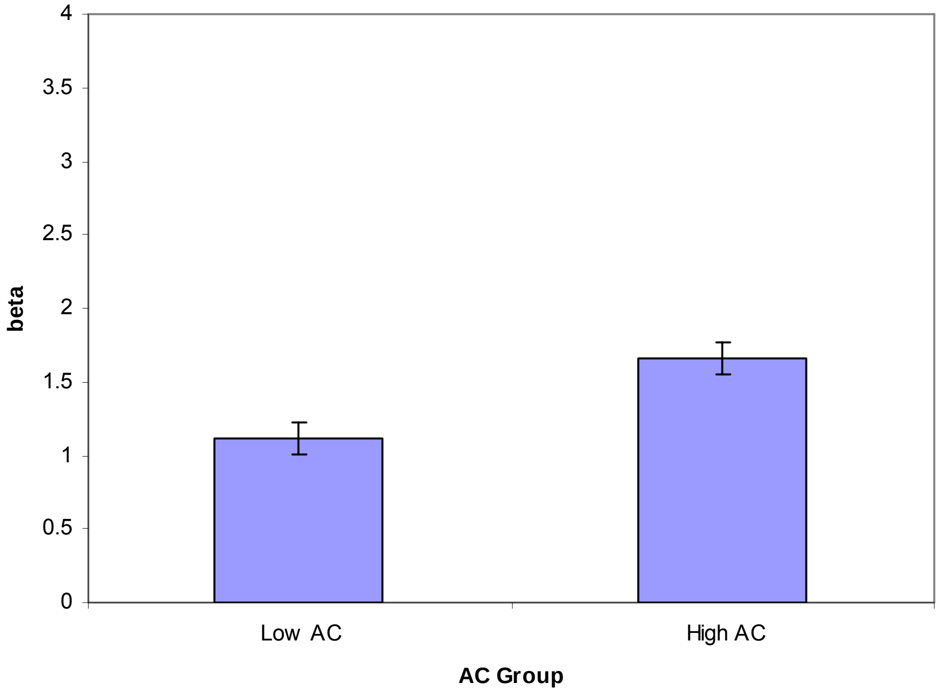


Figure 41. Mean machine-aided β as a function of AC group.

There was also a significant interaction between AC and modality, *F*(1, 28) = 10.351, $p < .005$ (see Figure 43). Individuals with high AC had higher β for the auditory modality than for the visual modality (*M* = 2.074, *SD* = 1.280, and *M* = 1.248, *SD* = .703, respectively), *t*(15) = 2.247, $p < .05$, *d* = .91. Individuals with low AC, however, had a higher β for the visual modality than for the auditory modality (*M* = 1.358, *SD* = .706, and *M* = .870, *SD* = .507, respectively), *t*(15) = 2.347, $p < .05$, *d* = .98. Also, in the auditory modality, those with high AC had a significantly higher mean β than those with low AC (*M* = 2.074, *SD* = 1.280, and *M* = .870, *SD* = .507, respectively), *t*(30) = 3.498, *p* = .001, *d* = 1.33.

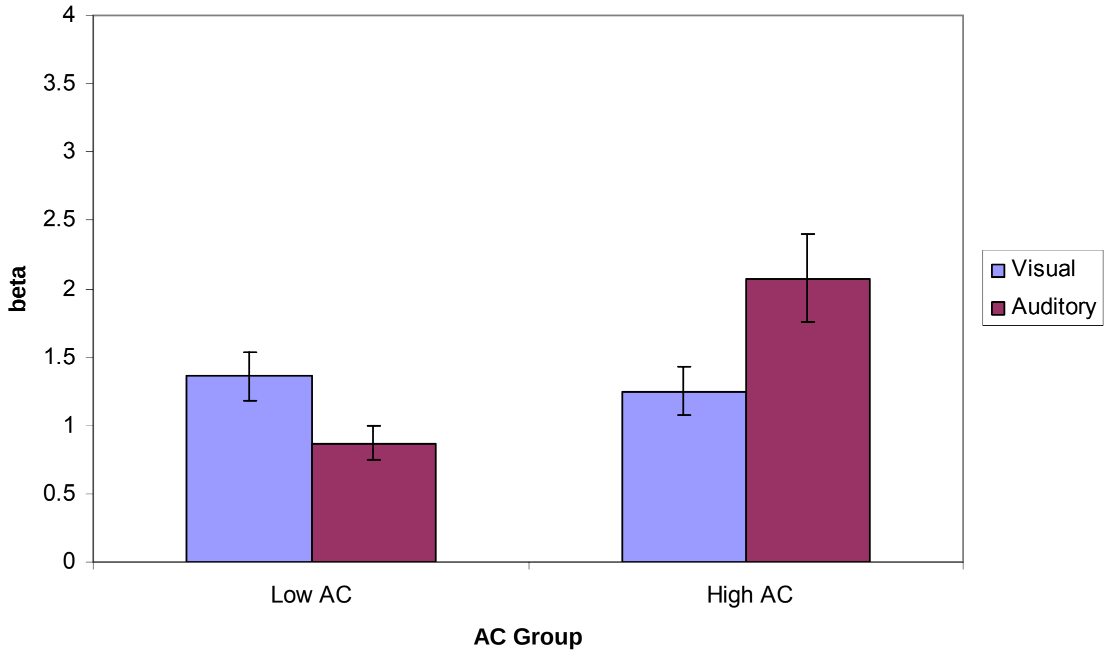


Figure 42. Mean machine-aided β by AC group and modality.

There was a marginal interaction among modality, load, AC, and LOA, *F*(1, 28) = 4.050, *p* = .054. There was a general trend for the high AC group to have a higher mean β in all tasks of both LOA groups, except the visual-combined task in LOA 4, in which low AC participants had a higher mean β than those in the high AC group. See Figures 44 and 45 for the machine-aided β by LOA per task for participants with low AC and high AC, respectively. Tests of simple effects were performed to compare the mean β in each task in each LOA within each AC group.

Those with low AC in LOA 2 had a higher mean β in the visual-only task than in the visual-combined task (*M* = 1.336, *SD* = .866, and *M* = .946, *SD* = .405, respectively), *t*(7) = 1.842, $p < .05$, *d* = .64. Also in LOA 2, the mean β of the auditory-combined task was higher than that of the auditory-only (*M* = 1.036, *SD* = 1.072, and *M* = .543, SD = .389, respectively), *t*(7) = 1.342. $p > .05$, *d* = = 1.79.

Those with low AC in LOA 4 had a higher mean β in the visual-combined task than on the visual-only task (*M* = 1.656, *SD* = .912, and *M* = 1.493, *SD* = .676), *t*(7) = 1.002, $p > .05$, *d* = .34. They also had a higher mean β in the auditory-only task than in the auditory combined task (*M* = 1.179, *SD* = .621, and *M* = .723, *SD* = .316, respectively), *t*(7) = 2.281, *p* = .057, *d* = 1.04.

For those with high AC in LOA 2, the visual-only β was higher than that of the visual-combined (*M* = 1.465, *SD* = .976, and *M* = 1.157, *SD* = .681, respectively), *t*(7) = .633, $p > .05$, *d* = .45. The auditory-only β was higher than that of the auditory-combined (*M* = 1.619, *SD* = 1.271, and *M* = 1.521, *SD* = 1.249), *t*(7) = .199, $p > .05$, *d* = .11.

Those with high AC in LOA 4 had a higher mean β in the visual-only task than the visual-combined task (*M* = 1.564, *SD* = 1.412, and *M* = .807, *SD* = .498, respectively), *t*(7) = 1.935, $p > .05$, *d* = .76. There was also a higher mean β in the auditory-combined task than in the auditory-only (*M* = 2.987, *SD* = 2.690, and *M* = 2.168, *SD* = 1.257, respectively), *t*(7) = .721 = $p > .05$, *d* =.92.

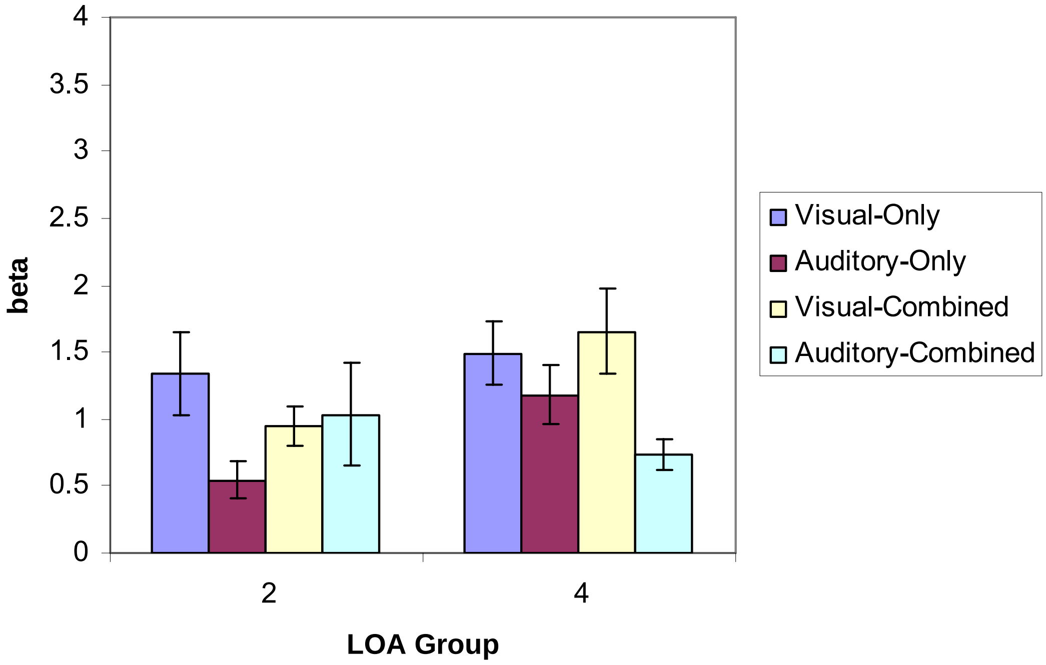


Figure 43. Mean machine-aided β by LOA, modality, and load for participants with low AC.

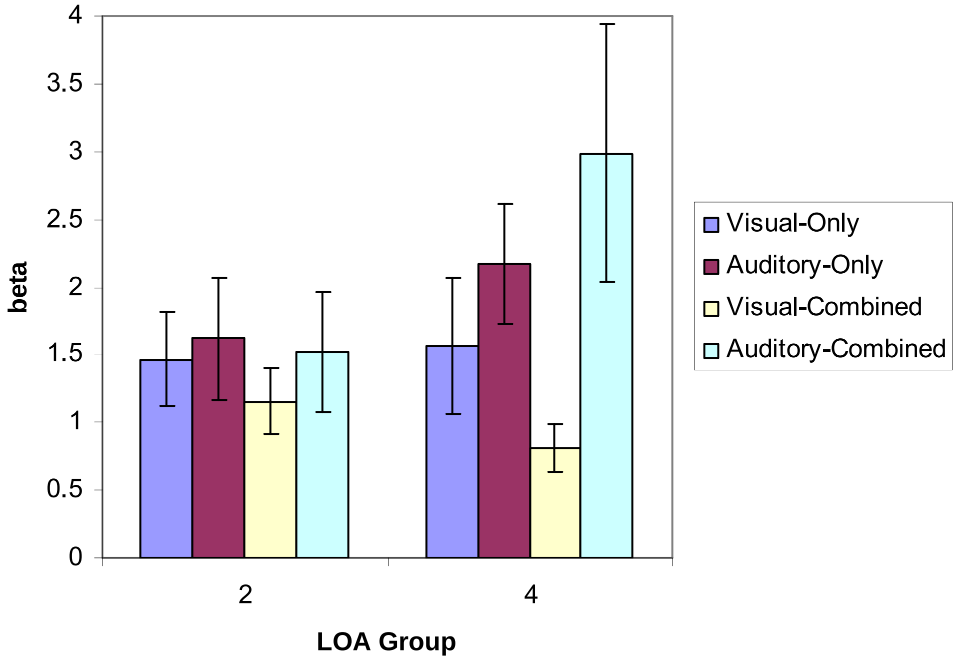


Figure 44. Mean machine-aided β by LOA, modality, and load for participants with high AC.

A separate set of a priori tests of simple effects was done to compare the operator-alone and machine-aided β in LOA 2. Operator-alone and machine-aided β were not significantly different (*M* = 1.272, *SD* = .565, and *M* = 1.203, *SD* = .536, respectively), *t*(15) = .412, *p* > .05, *d* = .13. Tests of simple effects comparing the tasks did not reveal any significant differences. In LOA 2, the machine-aided β was higher than the operator-alone β in the visual-only task (*M* = 1.400, *SD* = .894, and *M* = 1.323, *SD* = .596, respectively, *d* = .10). In all other pairs, there was a trend for operator-alone β to be higher than machine-aided β: the auditory-only task (*M* = 1.282, *SD* = 1.180, and *M* =

1.081, *SD* = 1.064, respectively, *d* = .18), the visual-combined task (*M* = 1.084, *SD* = .559, and *M* = 1.051, *SD* = .552, respectively, *d* = .06), and the auditory-combined task (*M* = 1.399, *SD* = 1.434, and *M* = 1.152, *SD* = .288, respectively, *d* = .24). See Figure 46 for a graph depicting the operator-alone and machine-aided β in each task in LOA 2.

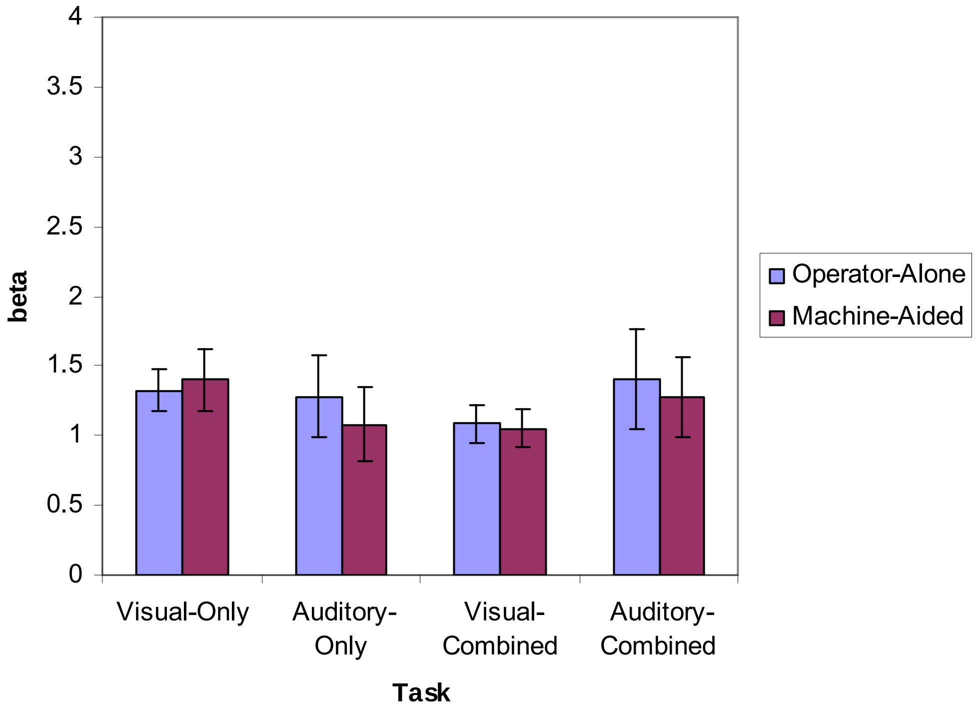


Figure 45. Mean operator-alone and machine-aided β in each task for LOA 2.

Also in LOA 2, there was a trend for the high AC group to have a higher mean β than the low AC group in all tasks and conditions of machine use (see Figure 47). Tests of simple effects showed that the machine-aided β of the high AC group was significantly higher than that of the low AC group (*M* = 1.619, *SD* = 1.271, and *M* = .543, *SD* = .389, respectively), *t*(14) = 2.288, *p* < .05, *d* = 1.20. All other pairs were nonsignificant (all *p* >

.05), however, the high AC group had a higher mean β in all tasks and conditions of machine use compared to the low AC group.

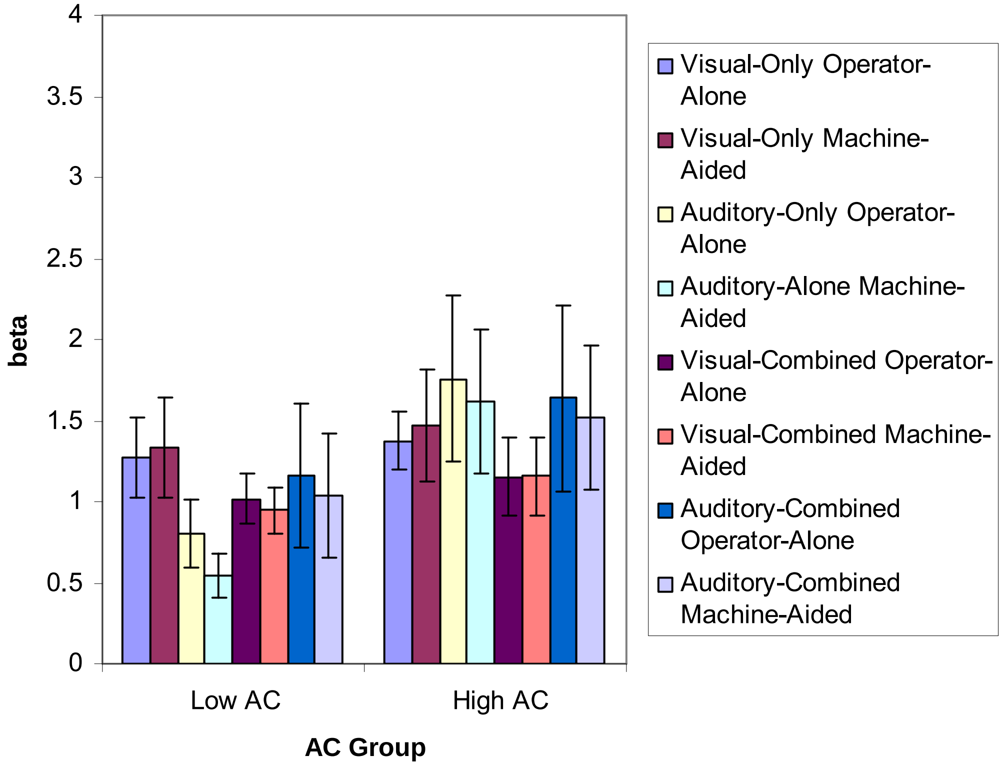


Figure 46. Mean operator-alone and machine-aided β per task and AC group in LOA 2.

See Figure 47 for mean β as a function of LOA group, modality, load, and machine use. This graph shows both operator-alone and machine-aided d' for both LOAs 2 and 4.

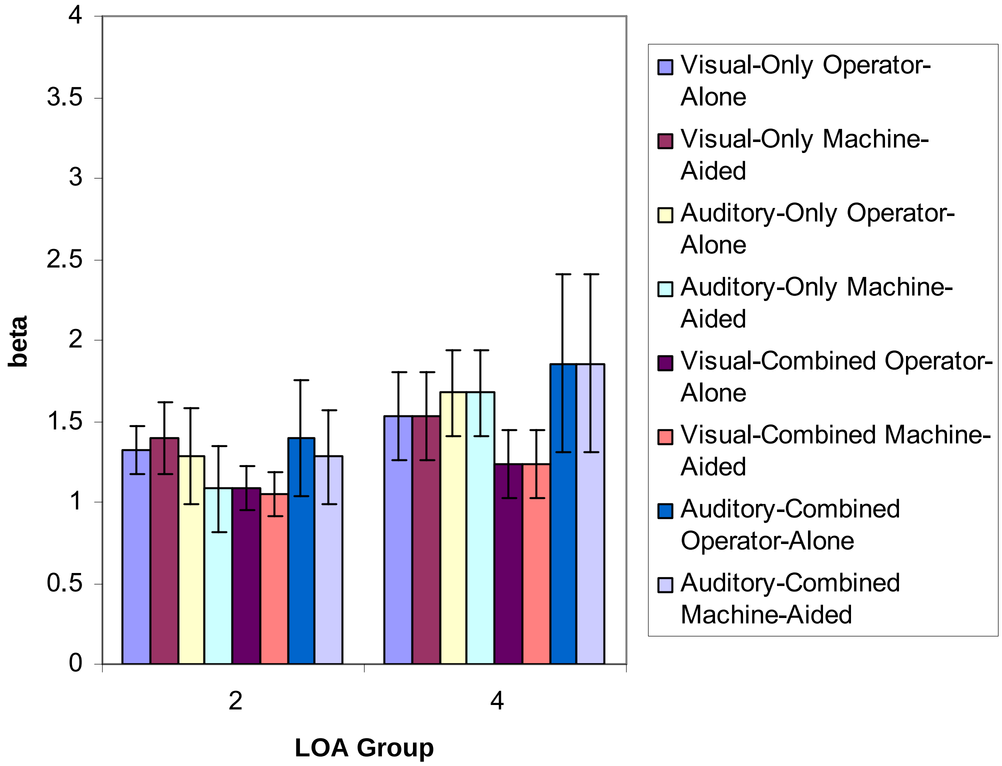


Figure 47. Mean β as a function of LOA group, modality, load, and machine use for LOA 2 and LOA 4.

Operator-Machine Interaction

Participants opted to agree with the machine response or veto it. There was a significant interaction between modality and LOA group, *F*(1, 28) = 27.254, *p* < .0001. In LOA 2, participants agreed with or vetoed the machine in the auditory modality more than in the visual modality (*M* = 23.500, *SD* = 15.218, and *M* = 17.219, *SD* = 11.454, respectively), *t*(15) = 2.950, *p* = .01, *d* = .47. The pattern was reversed in LOA 4, however, with participants responding to the machine more often in the visual modality than the auditory modality (*M* = 12.281, *SD* = 4.810, and *M* = 5.250, *SD* = 2.793, respectively), *t*(15) = 5.673, *p* < .0001, *d* = 1.79.

Figures 48 and 49 show the mean number of ‘agree’ and ‘veto’ responses for each task and in each LOA group for the low AC group and the high AC group, respectively.

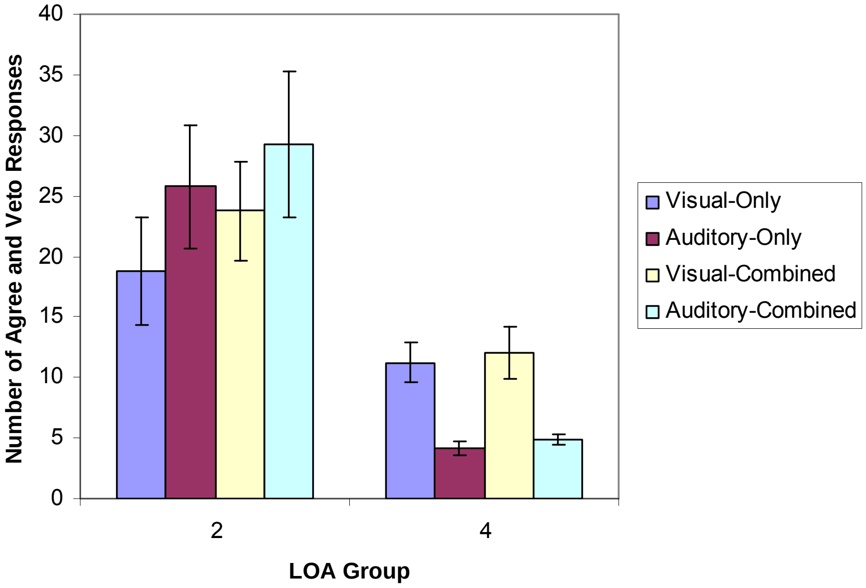


Figure 48. Number of ‘agree’ and ‘veto’ responses in each task for the low AC group.

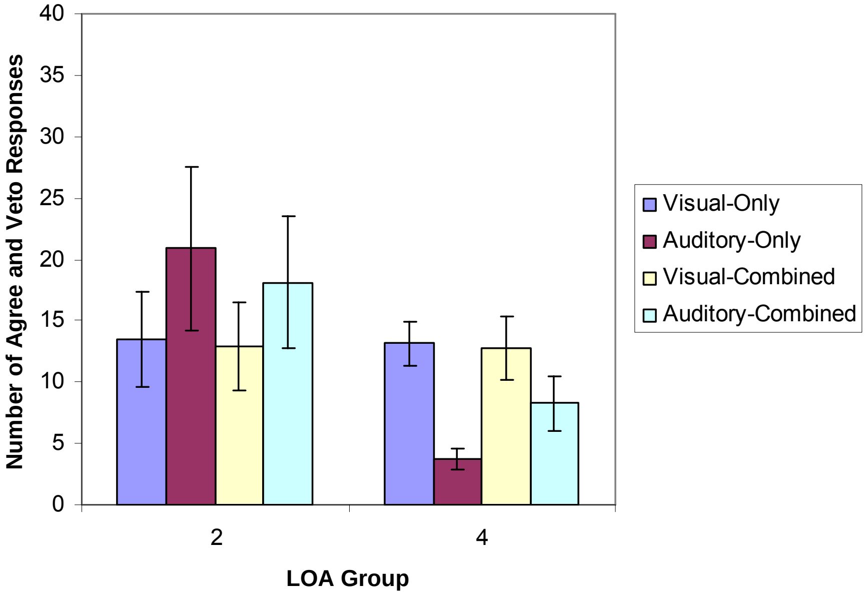


Figure 49. Number of 'agree' and 'veto' responses in each task for the high AC group.

Subjective Workload Ratings

As in Experiment 2, workload measurements were taken three times over the course of the experiment, once after each task: the visual-only, auditory-only, and combined. A 3-way (Task) within-subjects repeated measures ANOVA was used. AC and LOA were between-subjects variables. The main effects and interactions for the overall workload ratings can be seen in Table 12.

Table 12. Analysis of Variance for Workload Ratings

Source	*df*	*F*	η^2	*p*
		Between subjects		
Attentional Control (AC)	1	1.051	.036	.314
AC x *S* within-group error	28	(294.502)		
LOA	1	.143	.005	.708
LOA x *S* within-group error	28	(294.502)		
		Within subjects		
Task	2	4.559	.140	.042
Task x AC	2	.062	.002	.805
Task x LOA	2	.388	.014	.538
Task x AC x LOA	2	1.414	.048	.244
Task x *S* within-group error	28	(38.121)		

Note. Values enclosed in parentheses represent mean square errors. *S* = subjects.

* $p < .0001$

There was a significant main effect for task $F(2,28) = 4.559$, $p < .05$ (see Figure 51). The visual-only task had significantly higher workload ratings than the auditory-only task ($M = 45.177$, $SD = 16.919$, and $M = 40.406$, $SD = 17.392$, respectively, $t(31) = 2.558$, $p < .05$, $d = .28$. The combined task had a significantly higher mean workload rating than the visual-only task ($M = 48.516$, $SD = 19.256$, and $M = 45.177$, $SD = 16.919$, respectively, $t(31) = 2.175$, $p < .05$, $d = .18$. The combined task had a significantly higher mean workload rating than the auditory-only task ($M = 48.516$, $SD = 19.256$, and $M = 40.046$, $SD = 17.392$, respectively), $t(31) = 3.322$, $p < .005$, $d = .46$.

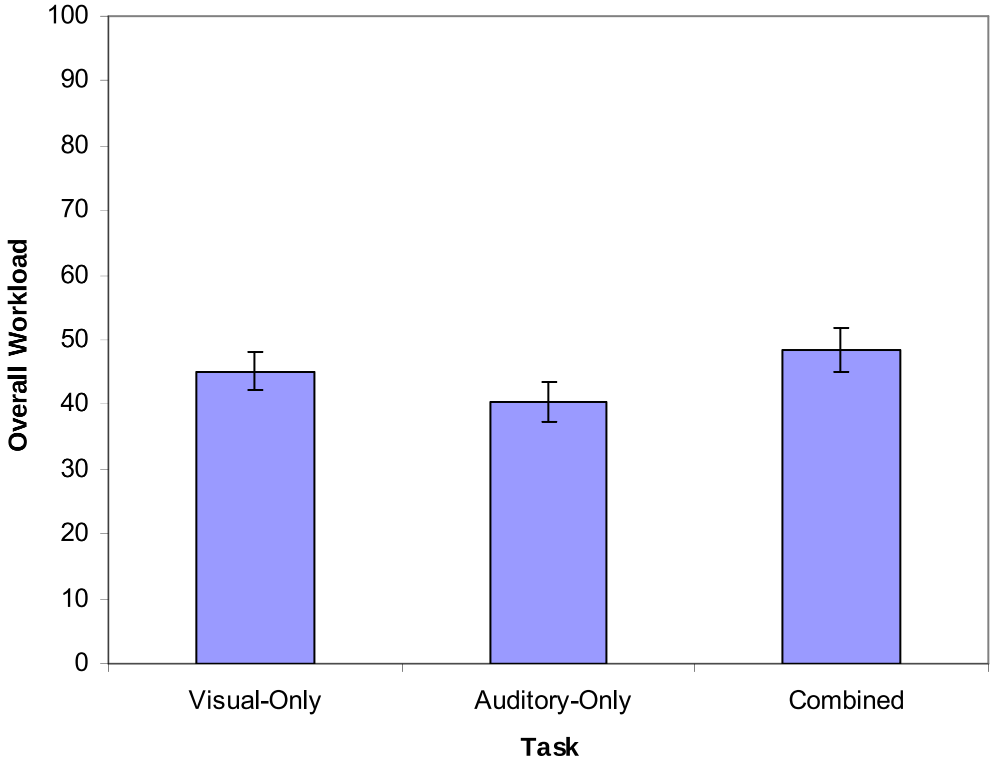


Figure 50. Overall workload by task.

There was no significant main effect for LOA, *F*(2, 28) = .143, *p* > .05. However, there was a trend for LOA 2 participants to report higher workloads across all three tasks (see Figure 52). LOA 2 participants reported higher workload that those in LOA 4 in the visual-only task (*M* = 46.000, *SD* = 19.728, and *M* = 44.354 *SD* = 14.176, respectively, *d* = .12), the auditory-only task (*M* = 41.229, *SD* = 17.146, *d* = .34), and combined task (*M* = 50.312, *SD* = 23.321, and *M* = 46.719, *SD* = 14.682, respectively, *d* = .22).

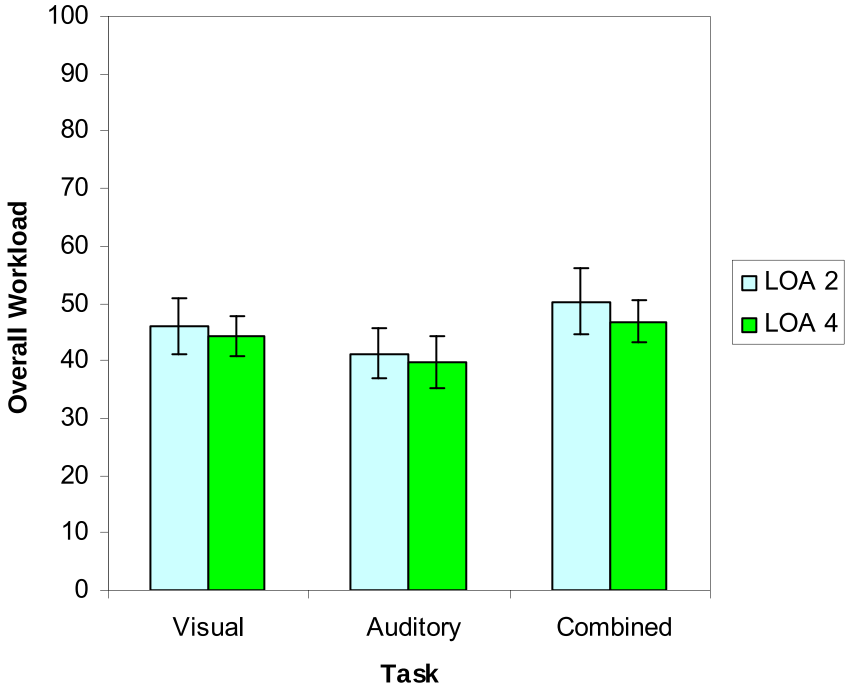


Figure 51. Mean overall workload for all three tasks as a function of LOA.

Although there was no significant main effect for AC, *F*(2, 56) = .090, *p* > .05, there was a trend for high AC participants to report higher workloads across all three tasks (see Figure 52). High AC participants reported higher workload that those with low AC in the visual-only task (*M* = 47.823, *SD* = 13.000, and *M* = 42.531, *SD* = 20.191, respectively, *d* = .58), the auditory-only task (*M* = 44.052, *SD* = 19.307, and *M* = 36.760, *SD* = 14.969, respectively, *d* = .53), and the combined task (*M* = 51.552, *SD* = 17.000, and *M* = 45.497, respectively, *d* = .12).

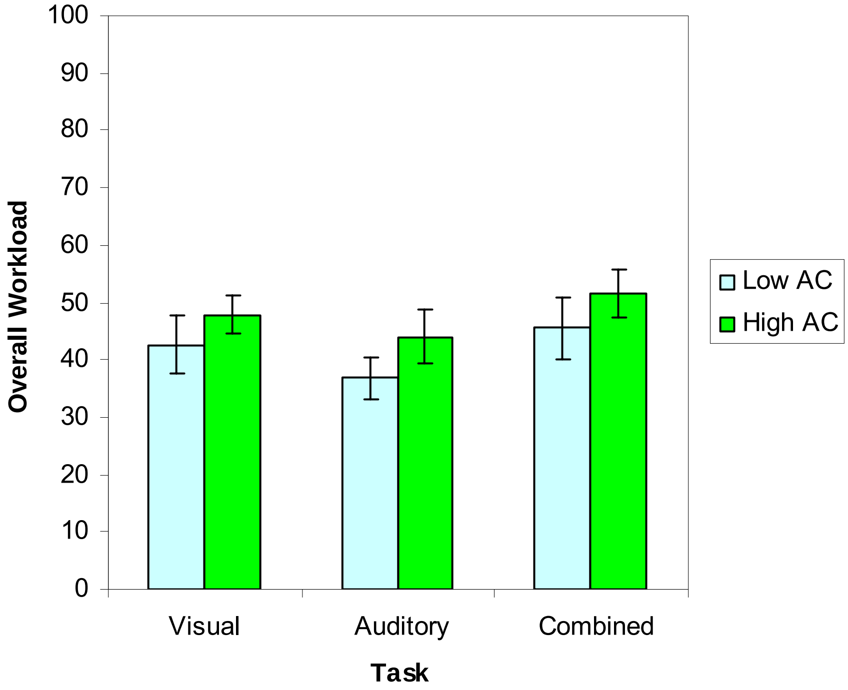


Figure 52. Mean overall workload for all three tasks as a function of AC.

Each subscale was assessed using a 3-way (Task) within-subjects repeated measures ANOVA. AC (see Figure 53) and LOA (see Figure 54) were between-subjects variables. See Table 13 for a summary of the significantly different workload scales across tasks.

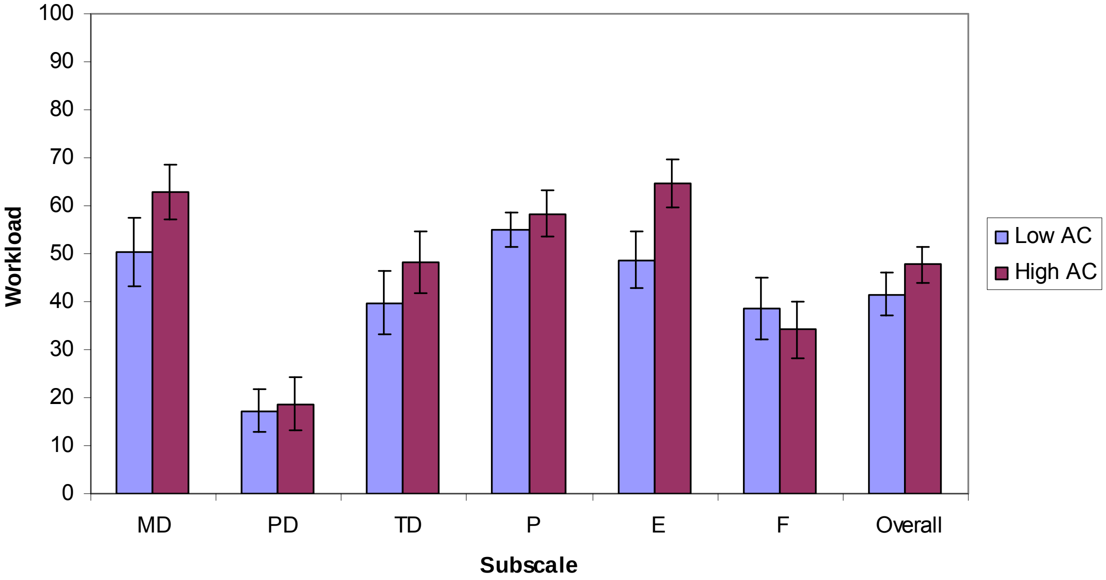


Figure 53. Mean workload across tasks for each subscale by AC group.

Note. MD = Mental Demand, PD = Physical Demand, TD = Temporal Demand, P = Performance, E = Effort, F = Frustration.

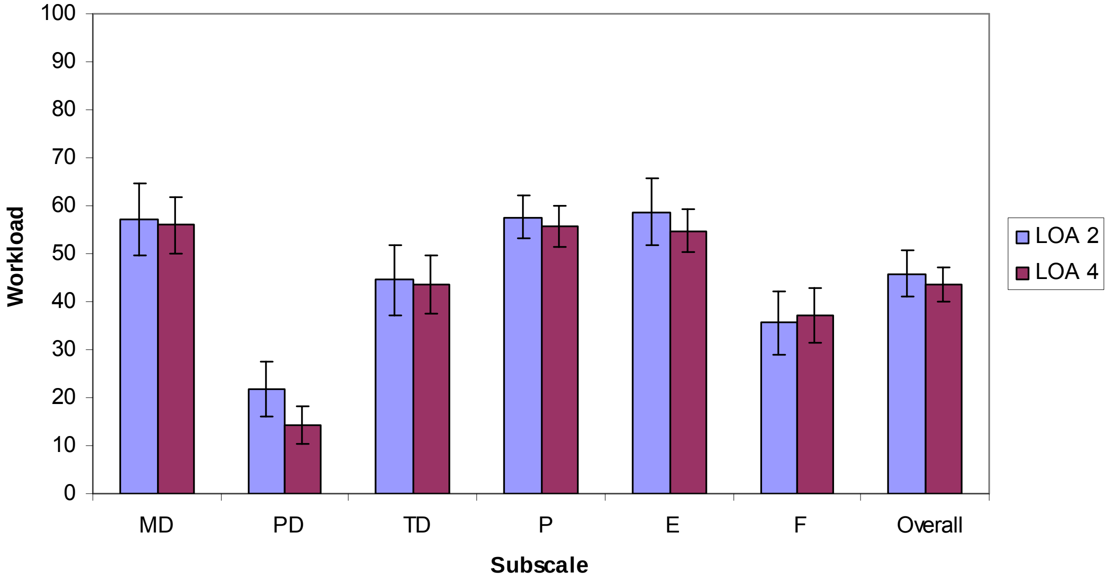


Figure 54. Mean workload for each subscale by LOA group.

Note. MD = Mental Demand, PD = Physical Demand, TD = Temporal Demand, P = Performance, E = Effort, F = Frustration.

For mental demand, there was a significant main effect for task $F(2, 56) = 15.018$, $p < .0001$. Mental demand in the visual-only task ($M = 61.188$, $SD = 27.413$) was higher than that of the auditory-only ($M = 43.31$, $SD = 28.020$), $t(31) = 4.138$. $p < .0001$, $d = .65$. Combined task mental demand ($M = 63.156$, $SD = 29.731$) was significantly higher than that of the auditory-only task, $t(31) = 5.233$, $p < .0001$, $d = .69$. There were no significant main effects for AC or LOA ($p > .05$).

For physical demand, there was a significant main effect for task $F(2, 56) = 4.217$, $p < .05$. Physical demand in the combined task ($M = 22.344$, $SD = 25.392$) was higher than that of the auditory-only task ($M = 14.688$, $SD = 20.433$), $t(31) = 2.771$, $p < .01$, $d = .33$. There were no significant main effects for AC or LOA ($p > .05$).

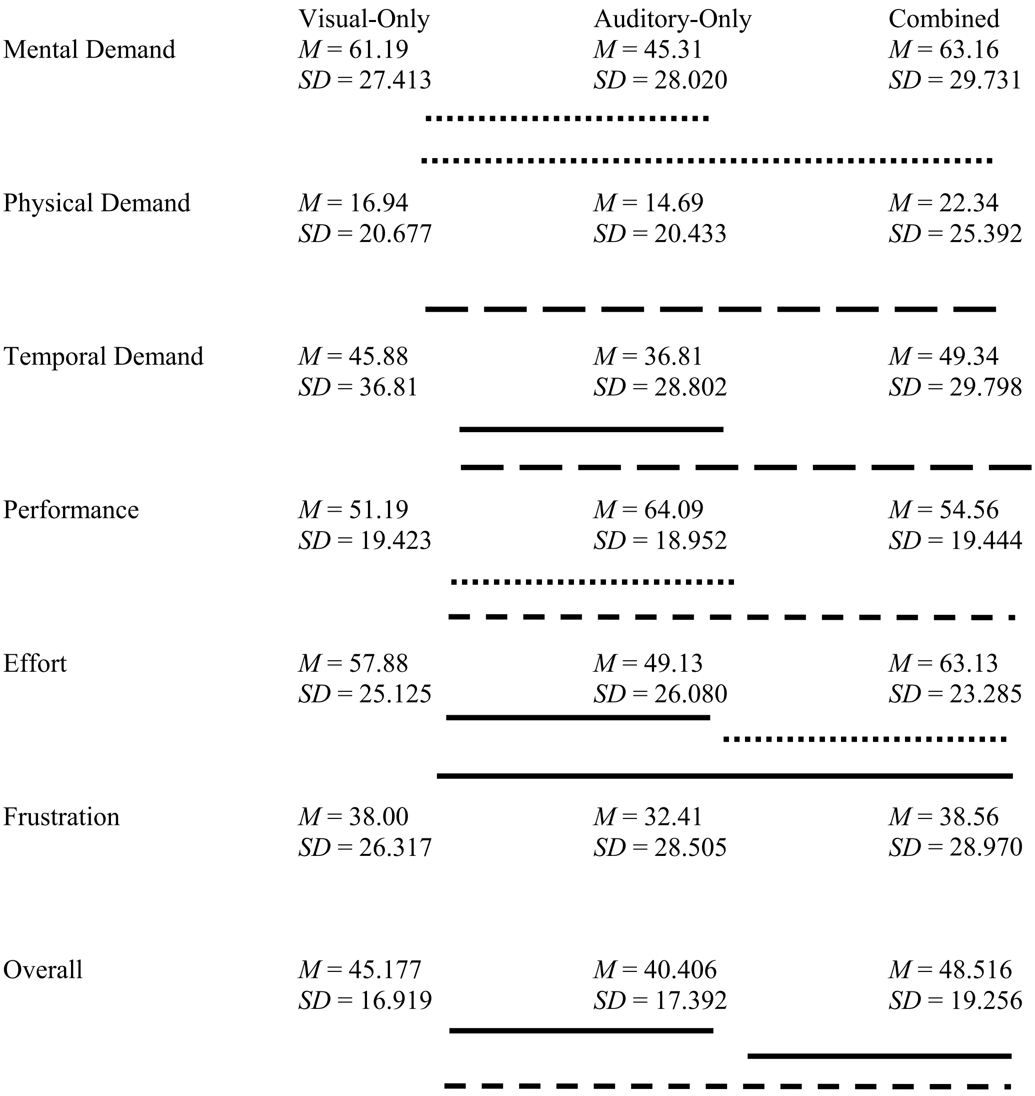
For temporal demand, there was a significant main effect for task *F*(2, 56) = 5.456, $p < .01$. Visual-only temporal demand (*M* = 45.875, *SD* = 28.971) was higher than that of the auditory-only (*M* = 36.813, *SD* = 28.802), *t*(31) = 2.677, $p < .05$, *d* = .31. Temporal demand in the combined task (*M* = 49.344, *SD* = 29.798) was significantly higher than that of the auditory-only task, *t*(31) = 2.815, $p < .01$, *d* = .43. There were no significant main effects for AC or LOA ($p > .05$). There was a marginal interaction between AC and LOA group, *F*(2, 28) = 3.228, *p* = .083. High AC participants in LOA 4 had higher temporal demand than those in LOA 2 (*M* = 55.96, *SD* = 26.506, and *M* = 40.500, *SD* = 32.507, respectively, *d* = .67). Low AC participants in LOA 2 had higher temporal demand than those in LOA 4 (*M* = 48.46, *SD* = 29.953, and *M* = 31.125, *SD* = 31.127, respectively, *d* = .83).

For performance, there was a significant main effect for task *F*(2, 56) = 11.981, $p < .0001$. Auditory performance (*M* = 64.094, *SD* = 18.952) was higher than that of the visual (*M* = 51.188, *SD* = 19.423), *t*(31) = 4.033, $p < .0001$, *d* = .67, and that of the combined task (*M* = 54.563, *SD* = 19.444), *t*(31) = 3.067, $p < .005$, *d* = .50. There were no significant main effects for AC or LOA ($p > .05$).

For effort, there was a significant main effect for task *F*(2, 56) = 11.300, $p < .0001$. Combined task effort (*M* = 63.125, *SD* = 23.285) was higher than that of the visual-only task (*M* = 57.875, *SD* = 25.125), *t*(31) = 2.418, $p < .05$, *d* = .22, and that of the auditory-only task (*M* = 49.125, *SD* = 26.080), *t*(31) = t(31) = 4.203, $p < .0001$, *d* = .57. Also, visual-only task effort was significantly higher than that of the auditory-only task *t*(31) = 2.668, $p < .05$. There was no significant main effect for LOA ($p > .05$). There was a marginal main effect for AC, *F*(2, 28) = 4.059, $p < .054$, such that those with high AC reported greater effort than those with low AC (*M* = 64.688, *SD* = 21.327, and *M* = 48.729, *SD* = 25.895, respectively, *d* = .87).

For frustration, there was no significant main effect for task *F*(2, 56) = 1.346, *p* > .05. There was a marginal interaction between task and LOA group, *F*(2, 56) = 3.036, *p* = .056. In the visual-only task, the frustration in LOA 2 was higher than in LOA 4 (M = 38.938, *SD* = 28.618, and M = 37.063, *SD* = 24.708, respectively, *d* = .09). In the auditory-only task, the frustration in LOA 4 was higher than in LOA 2 (M = 38.938, *SD* = 30.678, and, M = 25.88, *SD* = 25.437, respectively, *d* = .73). In the combined task, the frustration in LOA 2 was higher than in LOA 4 (M = 41.81, *SD* = 32.229), and M = 35.313, *SD* = 25.948, respectively, *d* = .28).

Table 13. Significant workload ratings by task.



	Visual-Only	Auditory-Only	Combined
Mental Demand	*M* = 61.19 *SD* = 27.413	*M* = 45.31 *SD* = 28.020	*M* = 63.16 *SD* = 29.731
Physical Demand	*M* = 16.94 *SD* = 20.677	*M* = 14.69 *SD* = 20.433	*M* = 22.34 *SD* = 25.392
Temporal Demand	*M* = 45.88 *SD* = 36.81	*M* = 36.81 *SD* = 28.802	*M* = 49.34 *SD* = 29.798
Performance	*M* = 51.19 *SD* = 19.423	*M* = 64.09 *SD* = 18.952	*M* = 54.56 *SD* = 19.444
Effort	*M* = 57.88 *SD* = 25.125	*M* = 49.13 *SD* = 26.080	*M* = 63.13 *SD* = 23.285
Frustration	*M* = 38.00 *SD* = 26.317	*M* = 32.41 *SD* = 28.505	*M* = 38.56 *SD* = 28.970
Overall	*M* = 45.177 *SD* = 16.919	*M* = 40.406 *SD* = 17.392	*M* = 48.516 *SD* = 19.256

Note. Significantly different pairs are indicated with lines. Each line patterns indicates a specific *p* value as follows:

——————— $p < .05$

— — — — - $p < .01$

- - - - - - - $p < .005$

········· $p < .0001$

Discussion

The primary goal of Experiment 3 was to assess AC effects upon operators in a fixed automation system rather than an adaptive automation system, as was manipulated in Experiment 2. This would allow for measurement of performance differences across the visual and auditory modality as well as single- and dual-task conditions when low and high AC operators were restricted to either a low or high LOA. The low LOA chosen for Experiment 3 was LOA 2, in which the operator retained the default responsibility of identifying all stimuli while immediately afterwards receiving an independent machine suggestion as to each stimulus identity; the operator could then change his or her original answer to match that of the machine if so desired. The high LOA chosen was LOA 4, in which the machine retained the default responsibility of identifying all stimuli, while the operator supervised its decisions and vetoed it when an error was suspected. As in Experiments 2 and 3, the machine was 90% reliable in its ability to correctly identify stimuli. Also as in Experiment 2, all participants completed a set of three temporal discrimination tasks: a visual-only, an auditory-only, and an audiovisual combined target detection task. It was hypothesized that d’ would be greater in LOA 4 than LOA 2, and subjective workload would be greater in LOA 2 than LOA 4, as operator workload and responsibility was lower in LOA 4. Also, as LOA 4 uses the machine decision-aid answers by default, it was predicted that the β in the high AC group would be more neutral in LOA 4 than in LOA 2, where the operator answers are the default.

Signal Detection Performance - d’ (Sensitivity)

As hypothesized, use of the higher LOA improved operator performance; the machine-aided d’ was higher in LOA 4 than the machine-aided d’ in LOA 2. Participants with both low and high AC benefited from using a higher degree of machine assistance.

In observing the d’ graphs from Experiments 2 and 3 (see Figure 55), there is an impression that the d’ values were generally higher in Experiment 3 wherein the LOAs were fixed. Also, the difference between the auditory d’ and visual d’ was even greater in Experiment 3 (visual d’ *M* = 1.243, *SD* = .516, and auditory *M* = 2.792, *SD* = .581) than in Experiment 2 (visual d’ *M* = 1.143, *SD* = .552, and auditory *M* = 2.163, *SD* = .572). This may indicate that participants adapted to the fixed LOAs in Experiment 3, and may have actually benefited from not having a choice. It is possible that the operators’ preferences and needs may not be strongly linked. It is also possible that participants were distracted by the option to switch LOAs in Experiment 2. Switching LOAs may have served as an additional task demand and operators may have allocated more attention to the flexible LOA feature more so than the actual perceptual discriminations.

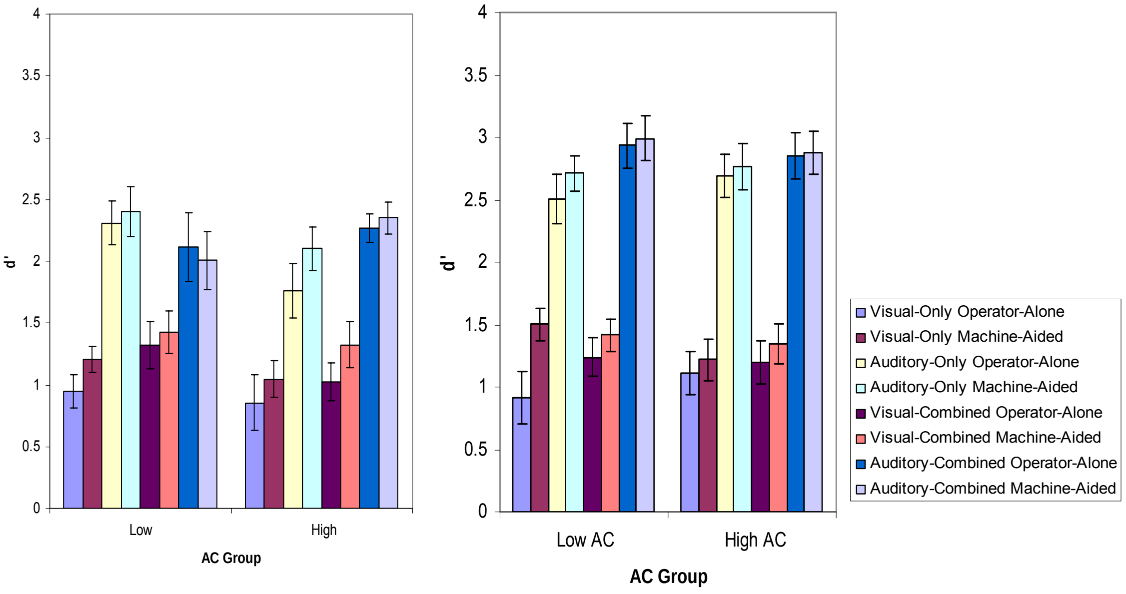


Figure 55. d' for each task and condition of machine use in Experiment 2 (left) and Experiment 3 across LOA 2 and 4 (right).

Based on the results of Experiment 2, a superior machine-aided d' in the auditory modality was also hypothesized, and this was supported. The difference was quite large; the auditory machine-aided d' was 2.836, while the visual machine-aided d' was much lower at 1.369. Thus, the modality effects persist under automation conditions where operators have both a choice of LOA and are fixed in either a low or high LOA. The modality effects are also consistently observed regardless of AC group, task load, and incorporation of the machine decision-aid.

In LOA 2, use of the machine aid improved the operators' d', indicating that the machine decision-aid was an effective tool as it was in Experiment 2. It significantly improved d' in the visual-only task; this corroborates previous findings that the visual tasks presented more of a challenge than their auditory counterparts and influenced participants to offload task responsibility to the machine. Across all tasks, there was a trend for the machine to improve the operator's d', and therefore the machine was helpful across modalities and loads. It also generally improved d' across AC groups.

There was no significant difference in d' between the low and high AC groups. This may be a task-dependent effect; the tasks employed in the experiments were monitoring tasks, and perhaps did not evoke enough higher-level processing and information integration from various potential sources. Thus, the high AC participants' advantage in focusing upon relevant information and shifting attention between tasks may not have been exploited.

Response Bias – β (Beta)

As in Experiment 2, the participants with high AC had a higher overall β than those with low AC. This suggests that β differences are retained given fixed LOAs as well as adaptive LOAs.

Also, as in Experiment 2, there was a significant interaction between AC and modality, such that the high AC group was especially conservative. There is thus a persistently strong tendency for those with high AC to perceive more stimuli as nons. At this point, the underlying mechanism may only be speculated. This may be driven by a tendency to perceive the shorter duration of time as longer. One future avenue of research might involve relationships between attentional control and time perception.

Experiment 3 presented an interesting interaction between AC and modality, such that the mean β for low AC participants was higher in the visual modality, but the mean β for the high AC participants was higher in the auditory modality. This effect was not observed in Experiment 2, wherein the modality effect was only present in the high AC group, who demonstrated a higher β for the auditory modality. This effect of higher visual β for the low AC group may have been a result of assigning participants to fixed LOA groups, and the low AC group may have been sensitive to the lack of choice. They may have compensated for this in Experiment 2 by choosing a suitable LOA to assist them in the visual tasks.

Specifically for individual tasks, the high AC group had a higher β in both LOAs and in all tasks except the visual-combined task completed at LOA 4. Here, the low AC group had the higher mean β. Also, in LOA 4, the high AC group had an extremely high β in the auditory-combined task (*M* = 2.987), which is interesting considering that the β values of LOA 4 should have been closer to 1, given the neutrality of the default machine answer. The high AC participants in LOA 4 were thus particularly conservative.

LOA 2 had both an operator-alone and machine-aided component, and so the mean β values of each were compared. In all tasks except for the visual-only task, the operator-alone β was higher than that of the machine aid; this was the case for both the low and high AC groups. In each task, the high AC group displayed a higher mean β than the low AC group. The trend for high AC participants to have a higher mean β therefore persists across conditions of machine decision-aid use as well, and was replicated across experiments and different task conditions. There may have been a task characteristic which inflated β across conditions; however, the low AC participants were regulated to a lower and thus more neutral β. However, the cause of this response bias is only subject to speculation at this point.

The tendency for the low AC group to maintain a more low-to-neutral β and the high AC group to maintain a higher β warrants future investigation. Because the instruction set did not associate the terms ‘target’ and ‘non’ with any sort of risk or threat, it is difficult to ascribe meaningful descriptions of any differences between low and high AC group motivations. The instruction set used in the current experiments differentiated ‘target’ from ‘non’ using a temporal discrimination, although Ohman (1996) has suggested that those with high TA have a low threshold for threat and may therefore interpret ambiguous stimuli as threatening. Future research may be able to further examine this association by manipulating the definitions of ‘target’ and ‘non’. One way to do this might be to assign a threat connotation to the definition of ‘target’, and then assess the possibility of threat-focus among those with different levels of AC and the inherently different levels of TA.

Subjective Workload Ratings

The patterns of means of each of the subscales for each task also closely replicated those of Experiment 2, such that the combined task typically had the highest workload ratings, and the visual-only task was typically associated with higher workload than the auditory-only task. The nature of each task was therefore subjectively similar across conditions of having a choice of LOA as well as being fixed at a particular LOA.

As hypothesized, task performance using LOA 2 was associated with higher overall workload ratings than using LOA 4. Thus, higher task loads and operator responsibility induced higher subjective workload ratings, and increased use of automation can be used to reduce the operator's workload as suggested in previous literature (Olson & Sarter, 2000; Endsley, 1999).

However, is also important to consider the attentional behavior of the operator. The results of Experiment 3 replicated those of Experiment 2, in which there was a trend for the high AC group to report higher overall workload ratings than the low AC group. The AC effects upon workload thus persisted once participants were fixed into a particular LOA as well as provided with adaptive automation. The high AC operators may have inherently exerted more attentional effort to complete the tasks and thus experienced increased workload.

Implications for Results

Assigning low and high AC groups to fixed low and high automation indicated that many of the effects seen in Experiment 2 persisted, such as the difference in d' for the visual and auditory modalities, the high AC group's higher β than that of the low AC group, and machine effects with regard to d' improvements and β regulation. Experiment 3 also demonstrated that subjective workload is elevated in response to task completion at lower LOAs, regardless of AC group. As the low and high AC groups had been assessed in environments with both low and high AC and yielded comparable d' values across conditions in Experiments 2 and 3, as well as demonstrated trends towards an inversely proportional relationship with LOA preference in Experiment 2, it was of interest to assess the same variables within the medium AC group in order to examine automation effects across the range of AC.

CHAPTER 7: EXPERIMENT 4 – MEDIUM LEVEL OF ATTENTION CONTROL IN ADAPTIVE AUTOMATION

In Experiment 2, both the extreme groups of AC were assessed: low and high AC. While some significant differences were found between the two groups, it was of interest to also assess the behavior and performance of the medium AC group. Doing so would facilitate capturing LOA effects along the entire range of AC in terms of LOA preference, performance, response bias, and subjective workload.

Hypotheses

Based on the original hypotheses generated by the literature review as well as the results of Experiments 2 and 3, the following a priori predictions were made.

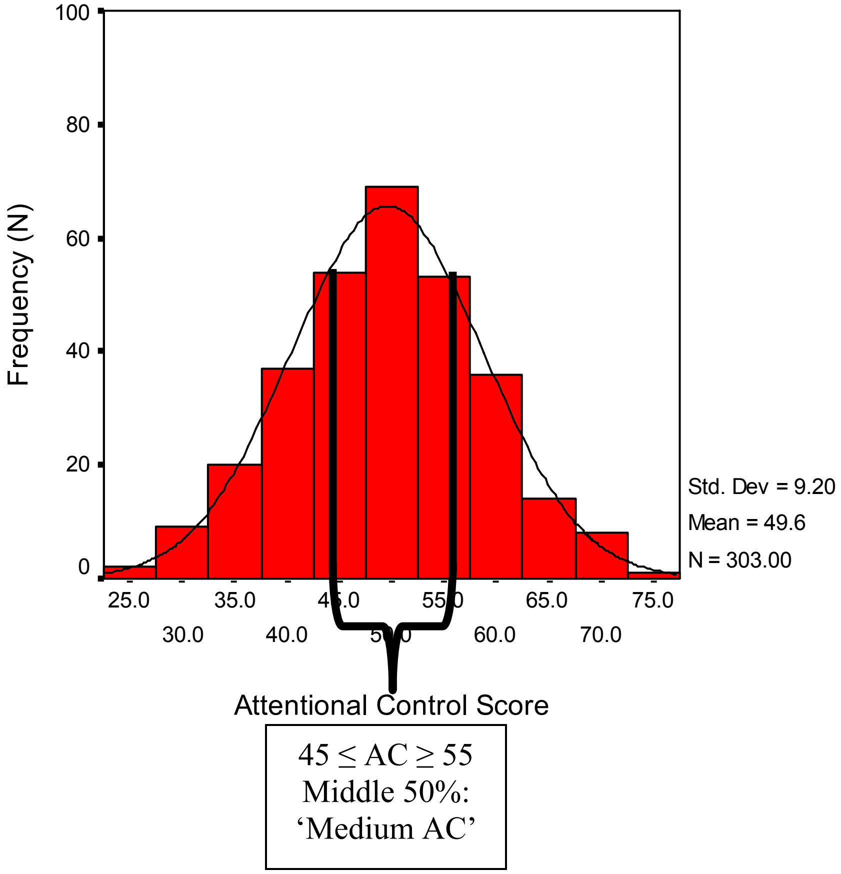
1. The LOA preferences will replicate those found in Experiment 2, such that statistically equivalent proportions of time will be spent at LOAs 1, 2, and 4, while significantly less proportion of time will be spent at LOA 3.
2. Participants will click on ‘agree’ and ‘veto’ more often in the auditory-only and auditory-combined tasks than in the visual-only and visual-combined tasks.
3. The d’ of the single-task conditions will be higher than that of the combined task.
4. Tasks drawing from the auditory modality will have a higher d’ than the visual modality.
5. Lower load will interact with the auditory modality to result in a higher d’.
6. The machine-aided β will be closer to 1 than operator-alone β. Thus, the machine-aid β will be lower than the operator-alone β.

7. The visual-only task will have higher workload ratings than the auditory-only task, and the combined task will result in higher workload ratings than the single-task conditions.

Methodology

Experimental Participants

Twelve participants were drawn from the pool generated in Experiment 1. Participants with medium AC scores (ranging from 45 to 55) were selected (see Figure 56). Ages ranged from 18 to 33 (*M* = 21.75, *SD* = 4.048). AC scores of selected participants ranged from 45 to 53 (*M* = 49.42, *SD* = 2.678). The participants consisted of two males and ten females. Among the male participants, the mean age was 27 (*SD* = 8.485 and the mean AC score was 49.5 (*SD* = .707). Among the female participants, the mean age was 20.7 (*SD* = 2.163 and the mean AC score was 49.4 (*SD* = 2.951).



45 ≤ AC ≥ 55
Middle 50%:
‘Medium AC’

Figure 56. Histogram which features the ‘Medium AC’ overlay upon the entire AC distribution. Participants were selected from the medium AC group.

Procedure

The procedure replicated that of Experiment 2, in which participants were trained to operate each of the four LOAs, and then change the LOA as often as they wished during the duration of the task.

As in Experiment 2, all participants completed the same visual-only, auditory-only, and audiovisual combined tasks, using the same stimuli types and response mechanisms for identifying targets and nons. They also completed the abbreviated version of the NASA-TLX after each of the three tasks.

Results

Preferences for LOAs

A 2 (Modality) x 2 (Load) x 4 (LOA) within-subjects repeated measures ANOVA was conducted.

See the ANOVA Table (Table 14) for a summary of the main effects and interactions for LOA preference in terms of proportion of time spent in each LOA.

Table 14. Analysis of Variance for LOA Preference (Time in LOA)

Source	*df*	*F*	η^2	*p*
		Within subjects		
Modality	1	1.000	.083	.339
Modality x Load	1	1.000	.083	.339
Modality x S within group error	11	(5.208E-07)		
Load	1	1.000	.083	.339
Load x S within group error	11	(5.208E-07)		
LOA	3	3.281*	.230	.033
LOA x Modality	1.334	1.781	.139	.205
LOA x Load	3	.556	.048	.648
LOA x Modality x Load	1.334	.247	.022	.695
LOA x S within group error	33	(.624)		

Note. Values enclosed in parentheses represent mean square errors. *S* = subjects.

*$p < .05$.

There was a significant main effect for LOA, $F(3, 33) = 3.281$, $p < .05$. A priori tests of simple effects indicated that more time was spent at LOA 2 than LOA 1 (M = .531, SD = .461, and M = .099, SD = .265, respectively), $t(11) = 2.511$, $p < .05$, $d = 1.15$. Significantly more time was also spent at LOA 2 than LOA 3 (M = .091, SD = .196), $t(11) = 2.614$, $p < .05$, $d = 1.24$). See Figure 57 for a pie chart depicting the proportion of time spent at each LOA across tasks.

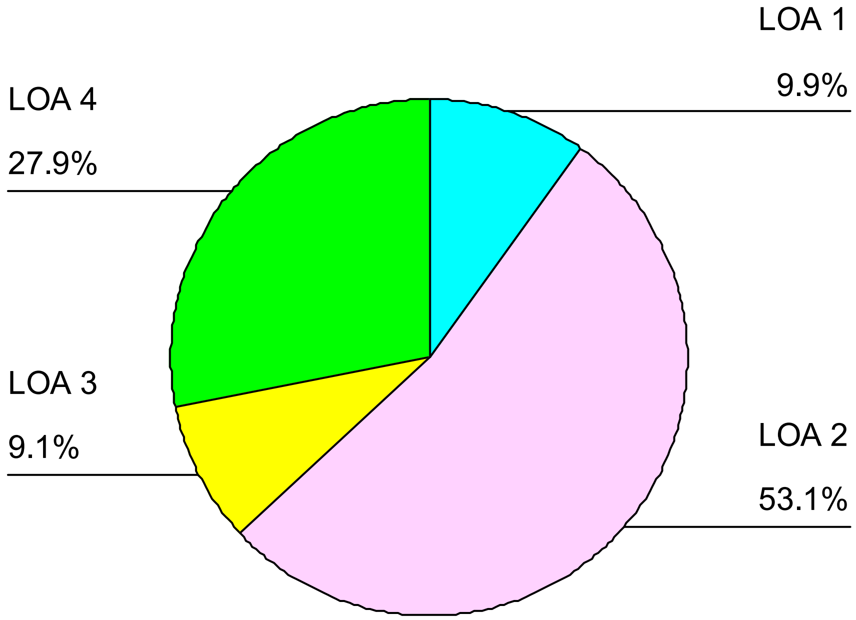


Figure 57. Proportion of time spent at each LOA collapsed across tasks.

There was no significant interaction between the initial LOA which was selected for each participant as the default LOA and the proportion of time spent at each LOA, *F*(3, 30) = 1.000, *p* > .05. This indicates that the once the task began, participants changed their LOA and they were not necessarily anchored to the system's default LOA.

There was no significant interaction between modality and LOA, *F*(3, 33) = .170, *p* > .05. See Figures 58 and 59 for depictions of proportion of time spent at each LOA in the visual modality and in the auditory modality, respectively.

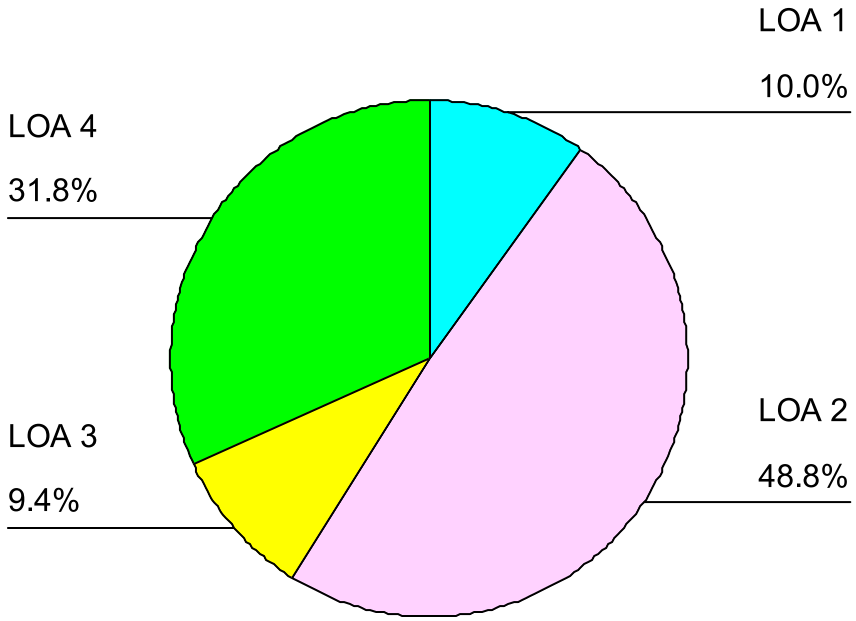


Figure 58. Proportion of time spent at each LOA in the visual modality.

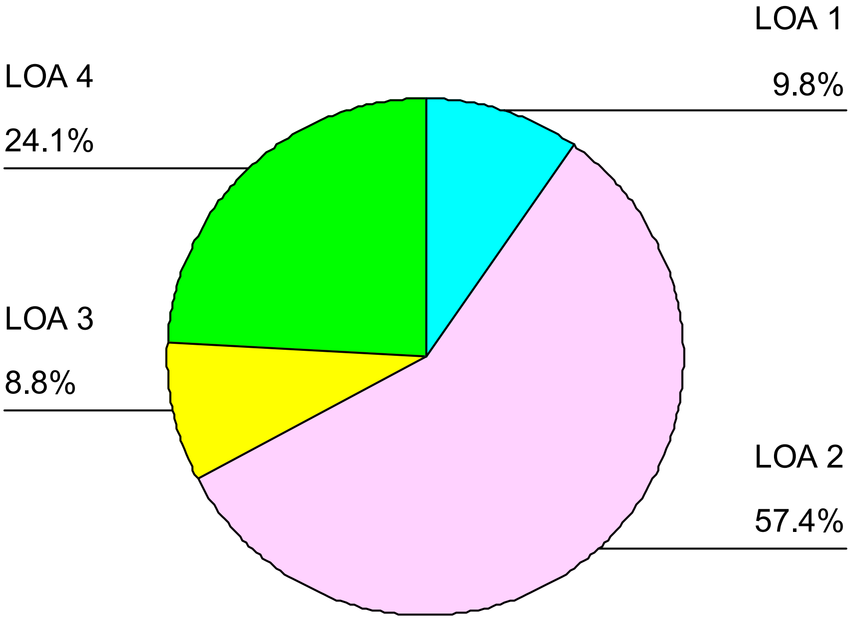


Figure 59. Proportion of time spent at each LOA in the auditory modality.

There was no significant interaction between LOA and load, *F*(3, 33) = .556, $p >$ .05

There was no significant interaction between LOA, load, and modality, *F*(1.334, 14.667) = .247, $p > .05$. Mauchley's test of sphericity was significant and thus the Greenhouse-Geisser adjustment for *df* was used. See Figures 60, 61, 62, and 63 for depictions of proportion of time spent at each LOA in the visual-only, auditory-only, visual-combined, and auditory-combined tasks, respectively.

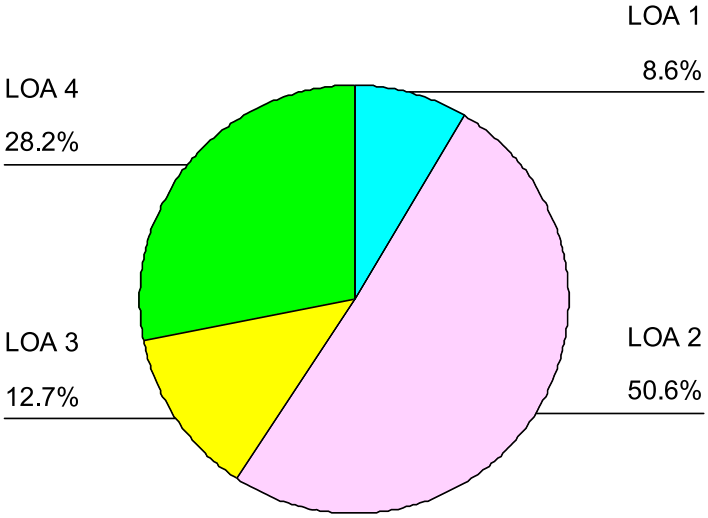


Figure 60. Proportion of time spent at each LOA in the visual-only task.

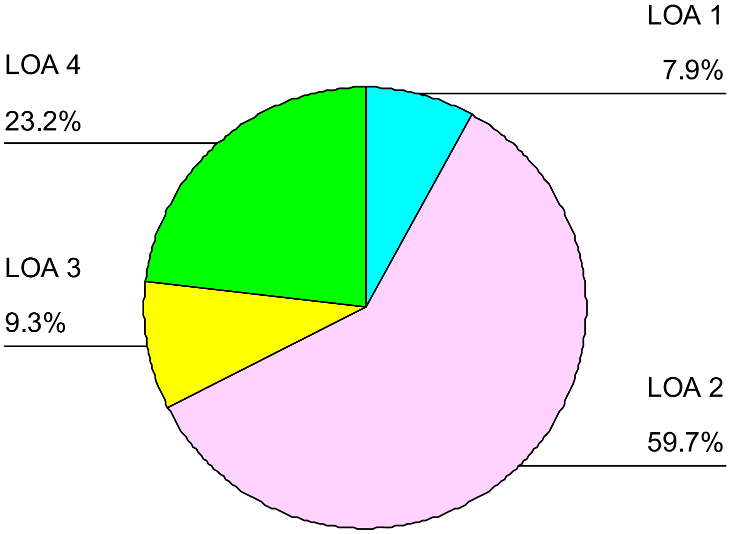


Figure 61. Proportion of time spent at each LOA in the auditory-only task.

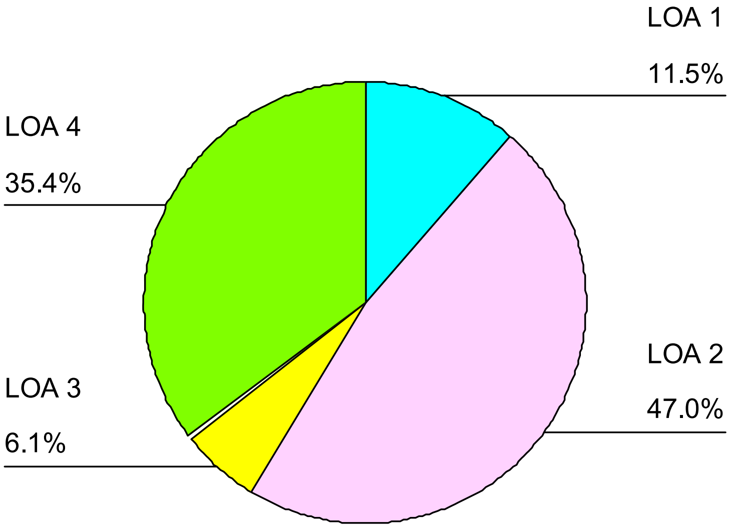


Figure 62. Proportion of time spent at each LOA in the visual-combined task.

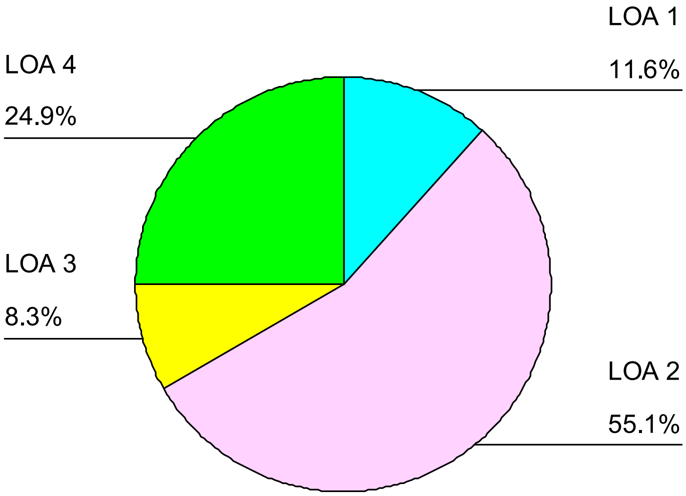


Figure 63. Proportion of time spent at each LOA in the auditory-combined task.

Operator-Machine Interaction

As in Experiment 2, participants were permitted to change LOAs according to their preferences. There was a marginal effect for load in terms of number of LOA changes, $F(1, 11) = 4.714$, $p = .053$, such that there were more frequent LOA changes in the single-task conditions than the combined task conditions ($M = 1.875$, $SD = 1.627$, and $M = 1.125$, $SD = .546$, respectively, $d = .62$).

In the visual-only task, the mean number of LOA changes was 1.92 (*SD* = 1.505). In the auditory-only task, the mean was 1.83 (*SD* = 1.749). In the visual-combined task, the mean was 1.08 (*SD* = .515). In the auditory-combined task, the mean was 1.17 (*SD* = .577). The Cohen's d for visual-only vs. visual-combined was .79, and the Cohen's d for auditory-only vs. auditory-combined was .53. See Figure 64 for a depiction of mean number of LOA changes by task.

The duty cycles for the visual-only, auditory-only, visual-combined, and auditory-combined tasks can be viewed in the appendices (Appendices W, X, Y, Z) of this document.

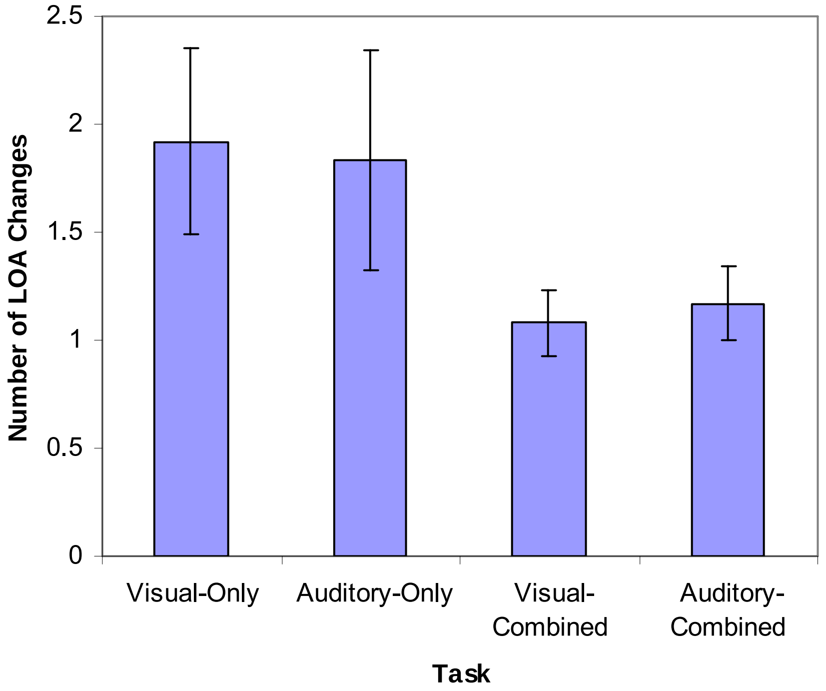


Figure 64. Mean number of LOA changes by task.

Also as in Experiment 2, participants opted to agree with the machine response or veto it. There was a significant main effect for modality, *F*(1, 11) = 9.675, *p* = .010, such that participants agreed with and vetoed the machine more often in the auditory modality than visual modality (*M* = 21.000, *SD* = 17.033, and *M* = 12.792, *SD* = 13.236, respectively, *d* = .53). There was no significant interaction between load and modality, *F*(1, 11) = .169, *p* > .05. There was a trend for participants to click on 'agree' and 'veto' more often in the auditory-only (*M* = 21.250, *SD* = 16.131) and auditory-combined tasks (*M* = 20.750, *SD* = 19.027, *d* = .03) than in the visual-only (*M* = 13.250, *SD* = 12.381), *d* = .56, and visual-combined tasks (*M* = 12.333, *SD* = 15.114), *d* = .57. Figure 65 depicts the number of agree and veto responses by each task.

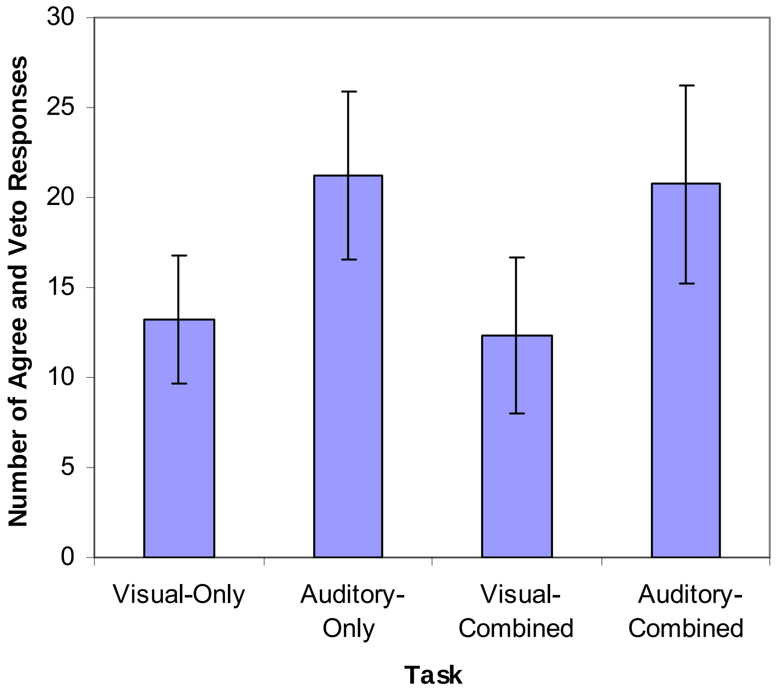


Figure 65. Mean number of times participants clicked on 'Agree' or 'Veto' during each task.

As in Experiment 2, during the single task condition, the operators were told that other task (the auditory or the visual) was fully automated, but that if they suspected a machine malfunction they should use the 'reset' button to recalibrate the machine. The 'reset' button was not pressed during the single task condition, however, during the combined task, one operator clicked on it twice, while another operator clicked on it three times.

Signal Detection Performance – d' (Sensitivity)

A 2 (Modality) x 2 (Load) x 2 (without-with Machine Aid) within subjects repeated measured ANOVA was conducted. The ANOVA table (see Table 15) summarizes the main effects and interactions involving d'.

Table 15. Analysis of Variance for d' (Sensitivity)

Source	*df*	*F*	η^2	*P*
		Within subjects		
Modality	1	116.367*	.914	.000
Modality x Load	1	.072	.007	.793
Modality x S within group error	11	(.268)		
Load	1	.337	.030	.573
Load x S within group error	11			
Machine use	1	.237	.021	.636
Machine x Modality	1	.134	.012	.721
Machine x Load	1	1.446	.116	.254
Machine x Load x Modality	1	.248	.022	.628
Machine x S within group error	11	(.123)		

Note. Values enclosed in parentheses represent mean square errors. *S* = subjects.

**p* < .0001.

There was a significant main effect for modality, *F*(1, 11) = 116.367, *p* < .0001, such that the auditory d' was higher than the visual d' (*M* = 2.471, *SD* = .591, and *M* =

1.331, *SD* = .424, respectively, *d* = 2.22). See Figure 66 for a depiction of the mean d' of each modality.

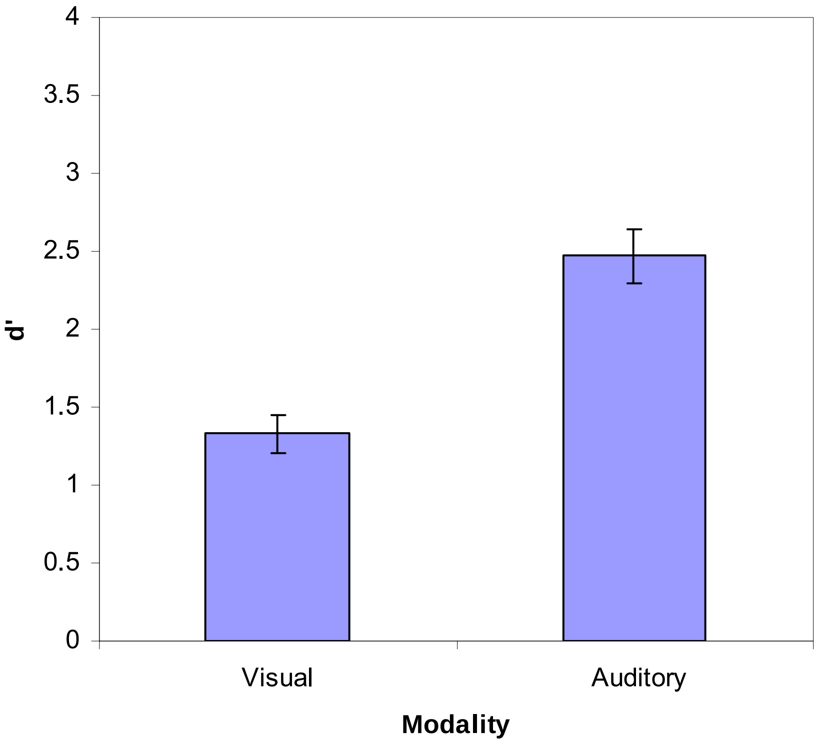


Figure 66. Mean d' for the visual and auditory modalities.

There was no main effect for load, *F*(1, 11) = .337, *p* > .05.

There was no significant main effect for machine use, *F*(1, 11) = .237, *p* > .05. There was no interaction between machine use and load, *F*(1, 11) = 1.446, *p* > .05. However, there was a trend for machine-aided d' to be higher than operator-alone d' in the single-task conditions, and operator-alone d' to be higher than machine-aided d' in the combined task conditions. The machine-aided d' was higher than the operator d' in the visual-only task (*M* = 1.467, *SD* = .906 and *M* = 1.215, *SD* = .710, respectively, *d* = .31), and in the auditory-only task (*M* =2.644, *SD* = .712 and *M* = 2.415, *SD* = 1.092, respectively, *d* = .25). The operator-alone d' was higher than the machine-aided d' in the visual-combined task (*M* = 1.458, *SD* = .496 and *M* = 1.186, *SD* = .830, respectively, *d* = .40), and in the auditory-combined task (*M* = 2.446, *SD* = .660 and *M* = 2.377, *SD* = .601, respectively, *d* = .11). Figure 67 depicts the mean operator-alone and machine-aided d' for each task.

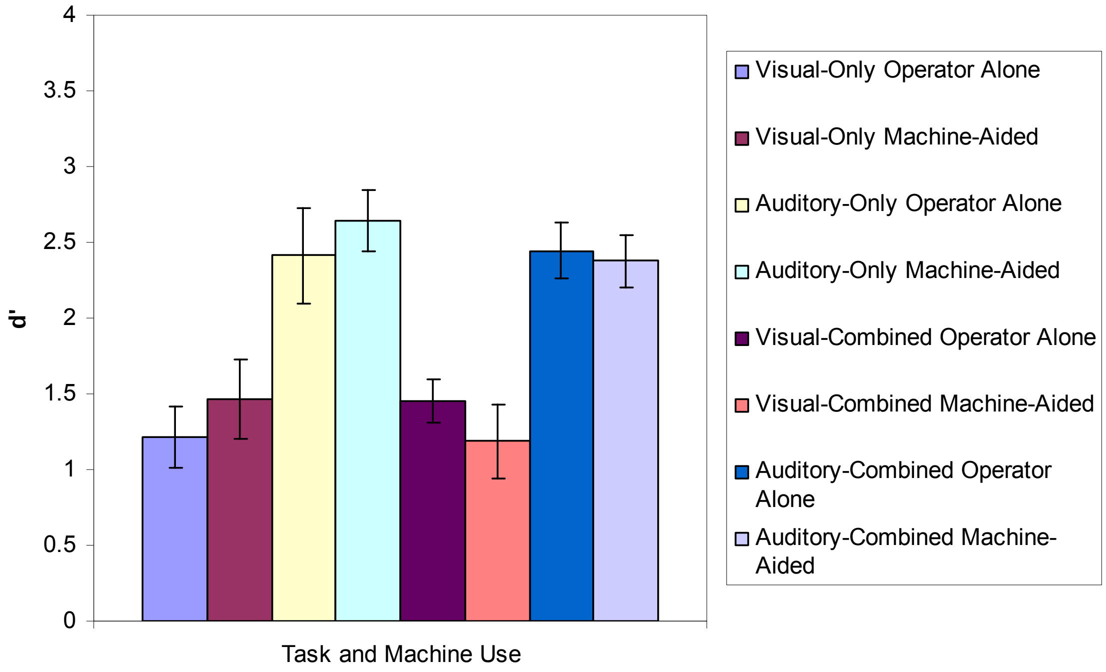


Figure 67. Mean operator-alone and machine-aided d' for each task.

Response Bias – β (Beta)

A 2 (Modality) x 2 (Load) x 2 (Machine Aid) within-subjects repeated measures ANOVA was conducted. The ANOVA table (see Table 16) summarizes the main effects and interactions involving β.

Table 16. Analysis of Variance for β (Beta)

Source	*Df*	*F*	η^2	*p*
		Within subjects		
Modality	1	.925	.078	.357
Modality x Load	1	.569	.049	.467
Modality x S within group error	11	(3.328)		
Load	1	.191	.017	.671
Load x S within group error	11	(1.275)		
Machine use	1	3.060	.218	.108
Machine x Modality	1	.400	.035	.540
Machine x Load	1	.588	.051	.459
Machine x Load x Modality	1	2.598	.191	.135
Machine x S within group error	11	(.347)		

Note. Values enclosed in parentheses represent mean square errors. *S* = subjects.

There were no significant main effects involving β (all $p > .05$).

There was a trend, however, for the machine-aided β to be lower than the operator-alone β in all four tasks: the visual-only task *(M* = 1.482, *SD* = 1.316 and *M* = 1.284, *SD* = .853, respectively, *d* = .10), the auditory-only task (*M* =1.538, *SD* = 1.589

and *M* = 1.518, *SD* = .1.508, respectively, *d* = .04), in the visual-combined task (*M* = 1.101, *SD* = .434 and *M* = 1.038, *SD* = .412, respectively, *d* = .15), and in the auditory-combined task (*M* = 1.920, *SD* = 1.548 and *M* = 1.361, *SD* = .870, respectively, *d* = .45). Figure 68 depicts the mean operator-alone and machine-aided β for each task.

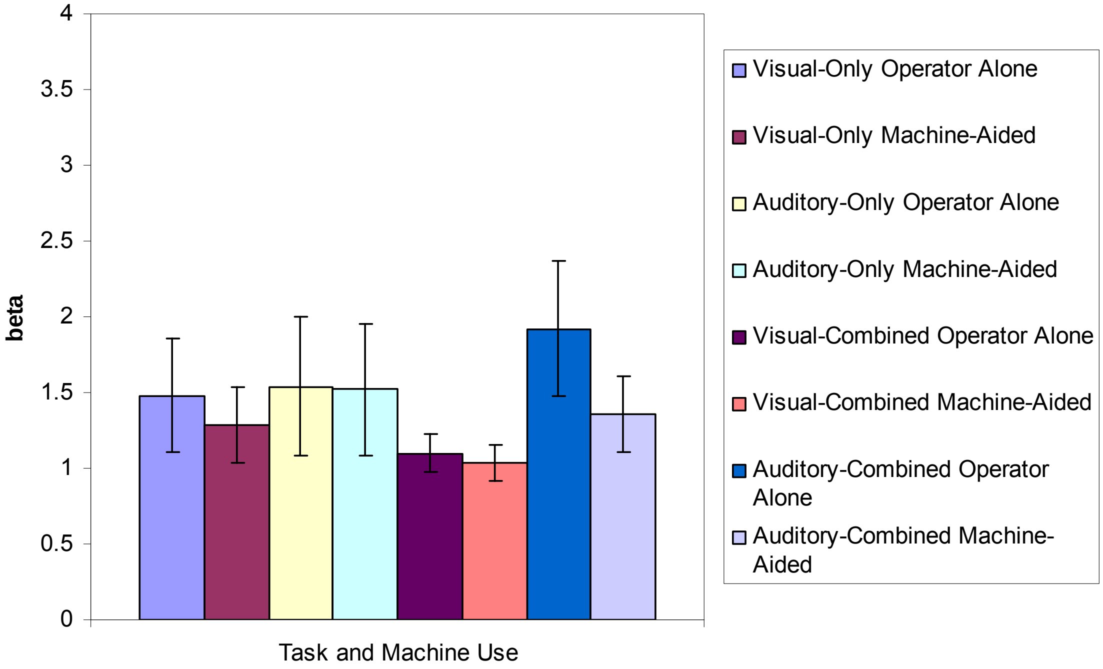


Figure 68. Mean operator-alone and machine-aided β for each task.

Subjective Workload Ratings

As in Experiment 2, workload measurements were taken three times over the course of the experiment, once after each task: the visual-only, auditory-only, and combined. A 3-way (Task) within-subjects repeated measures ANOVA was used.

For overall workload, there was a significant main effect for task *F*(2, 22) = 4.515, $p < .05$ (see Figure 69). A priori tests of simple effects showed that overall workload in the combined task was significantly higher than that of the auditory-only task (*M* = 56.264, *SD* = 12.038, and *M* = 47.806, *SD* = 12.775, respectively), *t*(11) = 3.754, *p* < .005, *d* = .68. Overall workload in the combined task was also higher than that of the visual-only task (*M* = 48.139, *SD* = 13.114), *t*(11) = 2.351, *p* < .05, *d* = .65.

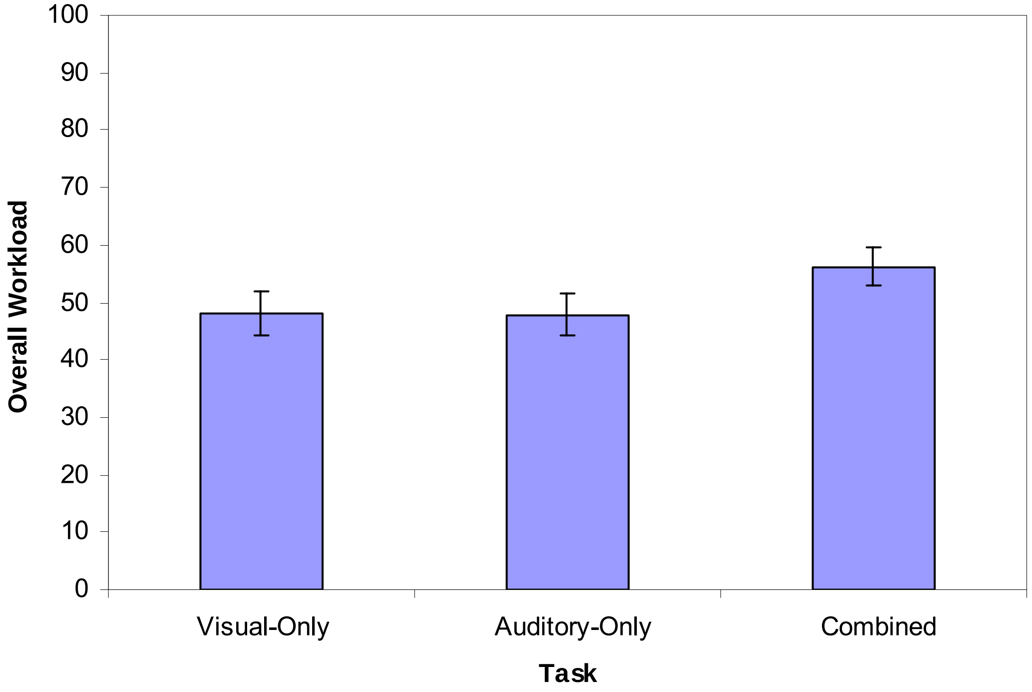


Figure 69. Overall workload as a function of task.

Each subscale was assessed using a 3-way (Task) within-subjects repeated measures ANOVA. A priori tests of simple effects were used to assess the significant differences among the tasks. See Table 17 for a summary of the significantly different workload scales across tasks. See Figure 70 for a graph depicting the subscale scores.

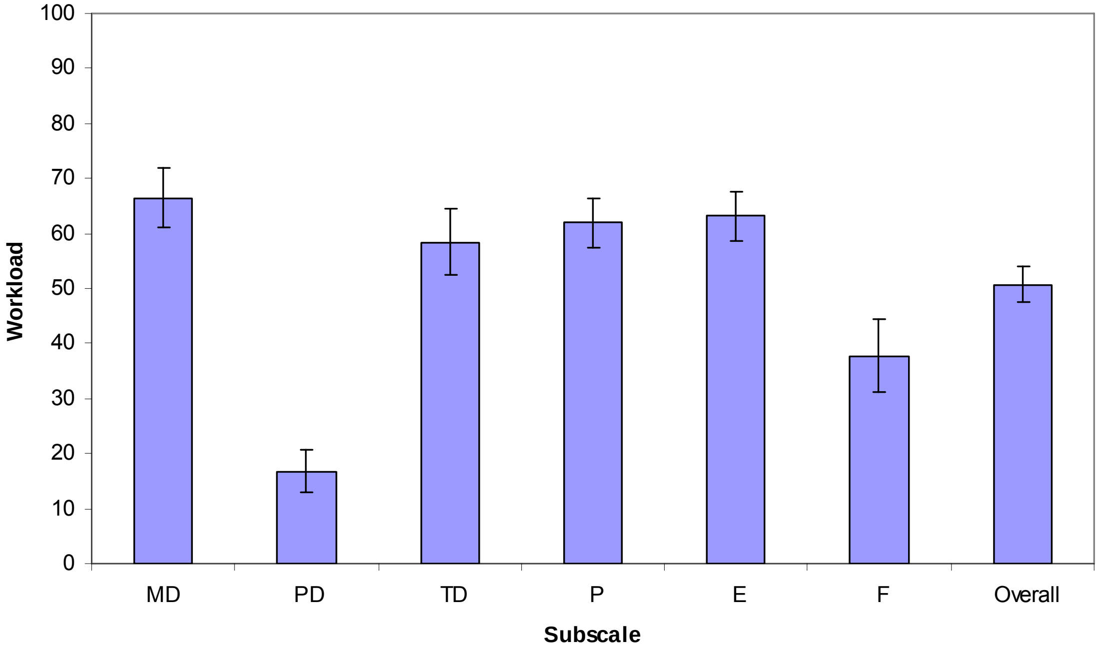


Figure 70. Mean workload for each subscale.

Note. MD = Mental Demand, PD = Physical Demand, TD = Temporal Demand, P = Performance, E = Effort, F = Frustration.

For mental demand, there was a significant main effect for task $F(2, 22) = 6.609$, $p < .01$. Combined task mental demand was significantly higher than that of the auditory-only task ($M = 76.25$, $SD = 14.943$, and $M = 55.00$, $SD = 25.965$, respectively), $t(11) = 4.659$, $p = .001$, $d = 1.00$.

For physical demand, there was a significant main effect for task $F(2, 22) = 5.272$, $p < .05$. Combined task physical demand was significantly higher than that of the auditory-only task ($M = 26.75$, $SD = 25.238$, and $M = 9.42$, $SD = 8.764$, respectively),

$t(11) = 2.542$, $p < .05$, $d = .92$. Combined task physical demand was also higher than that of the visual task ($M = 14.08$, $SD = 13.925$), $t(11) = 2.506$, $p < .05$, $d = .62$.

There were no significant main effects of task in the remainder of the workload subscales: temporal demand, $F(2, 22) = .386$, $p > .05$; performance, $F(2, 22) = 1.930$, $p > .05$; effort, $F(2, 22) = 2.139$, $p > .05$; and frustration, $F(2, 22) = 1.533$, $p > .05$.

Table 17. Significant workload ratings by task.

	Visual-Only	Auditory-Only	Combined
Mental Demand	*M* = 68.06 *SD* = 24.235	*M* = 55.00 *SD* = 25.965	*M* = 76.25 *SD* = 14.943
	- - - - -	- - - - - - - -	- - - - - -
Physical Demand	*M* = 14.08 *SD* = 13.925	*M* = 9.42 *SD* = 8.764	*M* = 26.75 *SD* = 25.238
		━━	━━━━━━
	━━━━━	━━━━━━━━	━━━━━━
Temporal Demand	*M* = 56.08 *SD* = 22.944	*M* = 58.75 *SD* = 25.424	*M* = 60.50 *SD* = 21.450
Performance	*M* = 58.00 *SD* = 19.088	*M* = 66.83 *SD* = 17.445	*M* = 60.83 *SD* = 18.678
Effort	*M* = 57.67 *SD* = 20.155	*M* = 61.25 *SD* = 21.701	*M* = 70.50 *SD* = 17.239
Frustration	*M* = 34.92 *SD* = 27.013	*M* = 35.58 *SD* = 21.732	*M* = 42.75 *SD* = 25.800
Overall	*M* = 48.14 *SD* = 13.114	*M* = 47.806 *SD* = 12.775	*M* = 56.264 *SD* = 12.038
		- -	- - - - - -
	━━━━━	━━━━━━━━	━━━━━━

Note. Significantly different pairs are indicated with lines. Each line patterns indicates a specific *p* value as follows:

━━━━━━━ $p < .05$

- - - - - - - $p < .005$

Discussion

The primary goal of Experiment 4 was to assess the preferences and performance of participants from the medium range of AC effects in an adaptive automation system, as was manipulated in Experiment 2. Four LOA choices were made available, ranging to fully manual operator control to mostly machine automated control in which the operator supervised the machine decisions and vetoed any suspected errors. As in Experiments 2 and 3, the machine was 90% reliable in its ability to correctly identify stimuli, with the intention of encouraging operator trust but discouraging full reliance. The experimental design used allowed for measurement of performance differences across the visual and auditory modality, as well as single- and dual-task conditions when medium AC operators were given a choice of LOA. Also as in Experiment 2, all participants completed a set of three temporal discrimination tasks: a visual-only, an auditory-only, and an audiovisual combined target detection task. It was hypothesized that the medium AC group would prefer medium LOAs, as Experiment 2 demonstrated a trend for the low AC group to use LOA 4 more often than the high AC group, and a trend for the high AC group to use LOA 1 more often than the low AC group. It was also hypothesized that the medium AC group would exhibit a mean β that intermediary to those of the low and high AC groups.

Preferences for LOAs

Medium AC participants preferred LOA 2 globally across tasks, and occasionally used LOA 3. Thus, the medium AC group preferred a moderately low LOA; they most often opted to retain the default task responsibility, but receive machine input as a suggestion. The medium AC group also spent more time at LOA 3 than both the low and high AC groups. The combined use of LOAs 2 and 3 suggests that the medium AC group was inclined to prefer the middle LOAs. LOA 1 was not frequently chosen. LOA 4 was chosen moderately and was the second most frequented LOA. This pattern of LOA preferences suggests that this group tended to prefer having at least some degree of machine assistance, whether in LOA 2, 3, or 4. Also, the medium AC group seemed to exhibit the same preference patterns across both the visual and auditory modalities and both loads and were thus consistent with their global preference of LOA 2. Figure 71 below graphically displays the proportion of time spent at each LOA across tasks for the medium AC group.

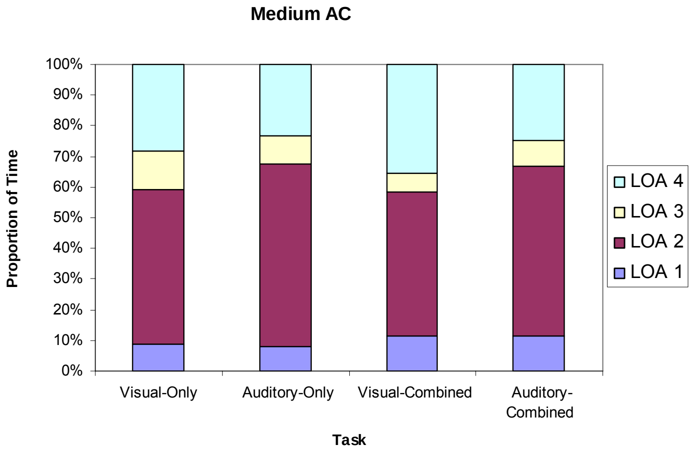


Figure 71. Proportion of time spent in each task for the medium AC group.

Sensitivity – d'

As with the low and high AC groups, the medium AC group demonstrated a significantly higher d' for the auditory modality than the visual modality. The higher sensitivity for the auditory modality thus persists despite degree of AC.

Effects of the machine decision-aid was differential between single- and dual-task conditions. The machine aid improved d' in the single tasks, however, it reduced d' in the dual-tasks. Although this interaction was not significant, it yields an interesting trend in which the mid AC group was overall less influenced by the machine aid in terms of d' than the low and high AC groups, both of which demonstrated d' increases after the machine answers were incorporated.

Response Bias – β (Beta)

The overall response bias of the medium AC group was slightly on the conservative side; however, the medium AC group was not as biased towards reporting nons as was the high AC group. However, similarly to the high AC group in Experiment 2, the medium AC group exhibited a high operator-alone β in the auditory-combined task, which was rendered more neutral after the machine-aid was incorporated.

The medium AC group demonstrated that their response bias was influenced by the machine decision-aid. There was a trend for the machine-aided β to be lower than the operator-alone β in all four tasks. Thus, the machine aid influenced the medium AC to have less of a response bias.

Subjective Workload Ratings

As with the low and high AC groups, the medium AC group reported the highest overall workload in the combined task. The mid AC group was thus sensitive to load in terms of subjective workload. Thus, the load of a dual-task condition is subjectively demanding to operators regardless of AC level. They also reported approximately that the overall workload ratings in the visual-only and auditory-only tasks were approximately equivalent, which is inconsistent with the overall lower d' in the visual-only task compared to the auditory-only task. This discrepancy suggests a potential dissociation between task difficulty and workload level, possibly due to an effect wherein performance is impacted by a specific task element, yet the subjective workload is reflecting a more global demand (Yeh & Wickens, 1988). The medium AC group may thus be sensitive to task characteristics other than those of the low and high AC groups, who reported higher visual-only workload and demonstrated a lower visual-only d' when compared to the auditory-only task.

Furthermore, in certain subscales (temporal demand, effort, and frustration), the auditory-only task had higher workload ratings than the visual only task. It appears that the medium AC group did not have a strong modality effect in terms of workload, and seemed to find the visual and auditory modalities equally demanding.

Implications for Results

The medium AC group did in fact demonstrate a preference for medium-range LOAs when compared to the low and high AC groups, indicating a trend for an inverse relationship between LOA preferences and AC. They also exhibited a mean β which appeared to be intermediary to the low and high AC groups. However, they also demonstrated a modality effect in terms of d’, but this appeared to be dissociated from their subjective workload ratings of the auditory-only and visual-only tasks. These results suggest that in many ways, the medium AC group performed differently than the extreme groups of AC, and may be governed by unique responses to task demand and automation use.

CHAPTER 8: GENERAL DISCUSSION

In this series of studies, participants with low, medium, and high AC levels were presented with an adaptive automation system, in which they could select the LOA they preferred to assist them in completing audiovisual temporal discrimination tasks. Low and high AC groups were also restricted to low and high LOAs in a fixed automation manipulation. Effects on perceptual sensitivity, response bias, and subjective workload were hypothesized to result from AC and LOA variations. The general findings are described here.

Preferences for LOAs

When participants within the middle range of AC scores were sampled, LOA preferences differed somewhat from the results of Experiment 2, in which low and high AC groups were sampled. Medium AC participants used LOA 2 to a greater degree than the low and high AC groups, and had a strong preference towards applying LOA 2 more than any other available LOA. They also used LOA 3 with greater frequency than the low and high AC groups. This trend indicates that medium AC participants were more inclined to use the middle LOAs. It can be inferred from Figure 72 that the low AC group gravitated towards LOA 4, the highest degree of automation, the medium AC group gravitated towards LOA 2 and even used LOA 3, wherein LOA 3 was largely neglected by the low and high AC groups. Finally, the graph indicates that the high AC group used LOA 1, the least amount of automation, more so than the low AC and mid AC groups. Those with higher AC thus had a preference for greater task control. This distribution of preferences across tasks supports the hypothesis that AC is inversely proportional to preferences for automation use, as low AC participants will use higher LOAs and high AC participants will use lower LOAs. This was complimented by a tendency for the medium range AC group to prefer the medium range LOAs as well. See Figure 72 for a comparison of how each AC group spent their time in terms of LOA across all tasks, and also see Figure 73 for a specific comparison for each of the four tasks.

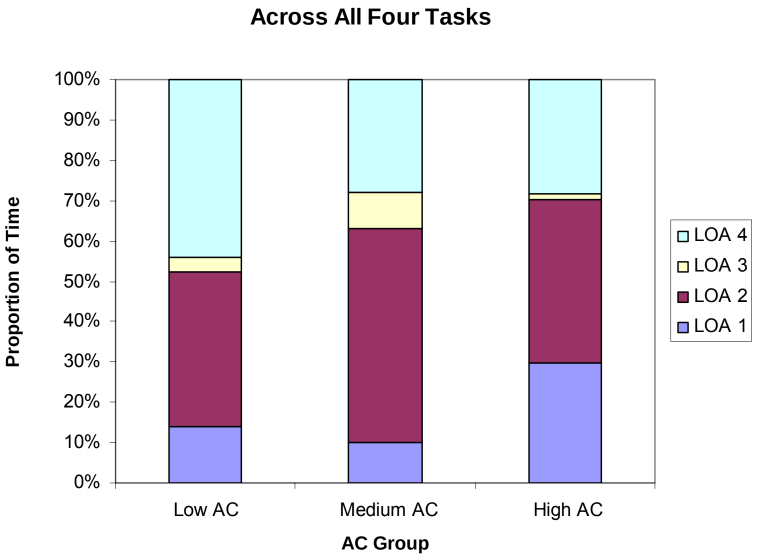


Figure 72. Proportion of time spent at each LOA for each AC group.

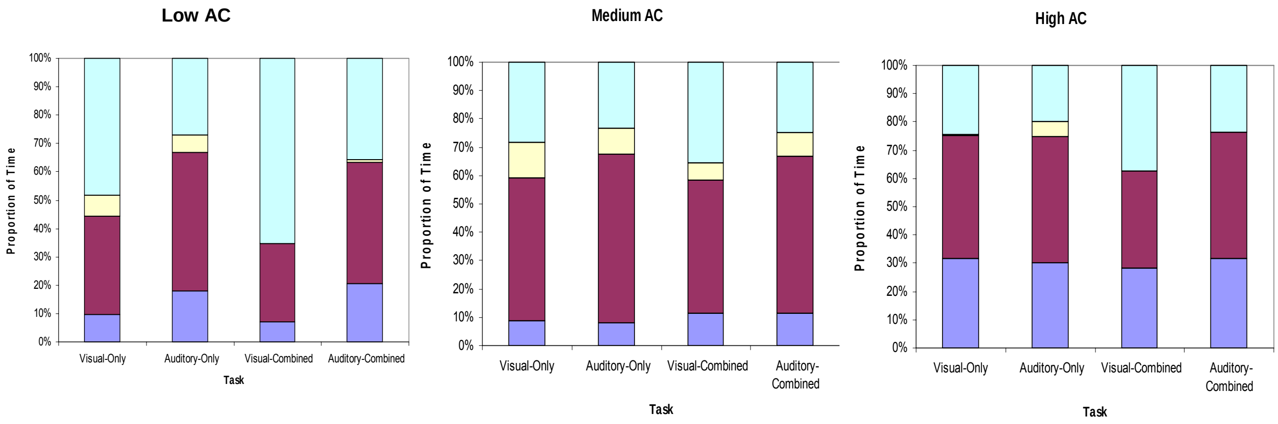


Figure 73. Proportion of time spent at each LOA in each task for each AC group.

LOA 2 was generally the most preferred LOA regardless of AC group. This preference endured despite operator traits and tasks. In LOA 2, participants first indicated their answer regarding the stimulus, and then received an independent suggestion from the machine. This response procedure thus gave the operator the primary responsibility; however, it also provided an opinion from a source which was 90% accurate. This machine opinion may have served as an appreciated confirmation to the operator’s first answer, and 90% reliability was considered adequate and still quite useful. The machine’s suggestion was therefore an important component of the operator’s decision-making, and operators preferred having a confirmation of their own answers.

The positive correlation between AC and DC found in Experiment 1 also supports these trends. Those with higher AC sought more task control, while those with low AC relinquished more control and preferred LOAs 2 and 4, while spending little time at LOA 1.

LOA 3 was generally not used often. The medium AC group used LOA 3 the most, however, this proportion of time was significantly less than the preferred LOA 2. LOA 3 thus continued to be largely unused across levels of AC, modality, and load. The reason for this may lie in the inherent nature of LOA 3 to integrate processing and responding strategies from both LOA 2 for stimuli in the left field and LOA 4 for stimuli in the right field. It is likely the case that LOA 3 demanded more effort in terms of developing strategies, thereby inducing higher subjective workload. Tulga and Sheridan (1980) state that the effort in developing strategies for task management can increase workload. Furthermore, Olson and Sarter (2000) state that automation can also increase workload if the operator is excessively taxed by interacting and using it. One way to assess this would be to fix a group of operators to each of the four possible LOAs and measure the consequent subjective workload reports associated with each LOA.

In Experiment 2, wherein low and high AC participants were sampled, LOA 4 was used in the dual-task conditions more often than in the single-task conditions. Thus, participants offloaded task demand to the machine when two sources of stimuli were presented. This was likely driven by the visual-combined task, in which both the low and high AC groups used LOA 4 the most often. In Experiment 4, however, this load effect did not emerge, although participants did use LOA 4 often in the visual-combined task. The effect of load may have thus been more dependent on the modalities of the tasks which comprised it; participants continued to use LOA 2 frequently in the auditory-combined task, but sought more machine assistance for the more challenging visual-combined task. The load effect may have thus been strongly associated with the nature of the visual discriminations in addition to the doubling of the event rate.

In Experiment 1, the mode response on the ALPI was to designate 75% of the control to the automation in the target detection task. This would correspond to LOA 4 in the task which was developed for Experiment 2, in which the machine had the default response and was 90% accurate, but the human operator could override it when a machine error was suspected. It is thus anticipated that LOA 4 will be a preferred LOA in Experiments 2 and 4. In Experiment 2, participants spent equally great proportions of time at LOAs 1, 2, and 4. Furthermore, in Experiment 4, participants spent the greatest proportion of time at LOA 2. This indicates that many participants felt that they would generally prefer more automation on the ALPI, but actually chose less automation in the experimental tasks. There may therefore be differences between the operator's hypothetical and practiced LOA preferences, which could be scenario and task-dependent.

Signal Detection Performance – d' (Sensitivity)

All three AC groups demonstrated higher mean d' values for the auditory tasks compared to the visual tasks. This greater saliency for auditory stimuli thus applies to operators regardless of AC characteristics. This corroborates previous research in which auditory signal detection has shown superiority over visual target detection in terms of accuracy (Davies & Parasuraman, 1982; Warm & Jerison, 1984; Szalma et al., 2004). The coupling effect may be the source of this discrepancy (see Hatfield & Loeb, 1968); wherein the auditory perception is advantageous because the source of the auditory stimulation is in close contact with the ears. However, visual processing requires that the eyes are focused upon the source of stimulation, necessitating selective attention to scan for information (Wickens & Hollands, 2000). Thus, while auditory stimuli and auditory perception are closely coupled, there is a looser coupling between visual stimulation and visual perception. Auditory stimuli are therefore more intrusive and difficult to neglect. Gugerty abd Tirre (1997) have found that situation awareness is associated with visual and temporal processing; the visual task may not have thus facilitated situation awareness due to the coupling effect, resulting in reduced visual d'. Task specificity driven by modality differences is therefore an important consideration in interpreting the experimental results.

The coupling effect should thus be considered in complex human-machine systems; the inherently elevated saliency of auditory stimuli should be assessed for its ability to both increase operator awareness of system status as well as override the saliency of competing visual stimuli. Thus, while multi-modal displays could potentially aid the human operator, as suggested by Sarter (2000) and Billings (1997), the results of the current study highlights the potential for visual warnings to be neglected in favor of auditory warnings. Adaptive automation, however, can in part ameliorate this effect by allowing the operator to receive machine aid in monitoring the less salient stimuli. The results of these studies did indeed demonstrate that use of the machine aid improved visual task performance. Further investigation, however, should assess the parameters of visual and auditory stimuli which are more psychophysically equated in complex systems.

The d' values for both the auditory and visual modalities were generally higher in Experiment 3, wherein operators were fixed in either LOA 2 or LOA 4, than in Experiment 2, wherein operators had a choice of LOA. Despite being allowed to choose any LOA, operators performed better in the situations in which the automation was fixed. This may have been due to a potential extra task demand of attending to LOA choices and switching LOAs in Experiment 2. Additionally, participants may have preferred a certain style of interaction associated with a particular LOA, but, this may not have necessarily been the most effective style. In Experiment 3, participants may have adapted to the LOA to which they were assigned.

AC effects upon d' were generally not observed in the perceptual tasks. According to Engle (2002), working memory capacity may be related to performance on higher-order tasks, where individual differences in attentional control facilitate maintenance of information in an active state. The perceptual tasks employed in Experiments 2, 3, and 4 may not have been taxing enough for individual differences in attentional control to emerge. AC effects were expected to emerge especially within the dual-task conditions; however, these effects were not observed. One possible reason for this is the use of two different modalities in the audiovisual combined task. According to Baddeley (1995), working memory involves a phonological store for representing information as words and sounds, as well as a visuospatial sketchpad for information such as visual images. With two separate stores, the visual and verbal codes do not compete for the same limited processing resources, allowing for time-sharing in a dual-task condition demanding both verbal-audio and spatial requirements. Thus, two concurrent visual tasks or two concurrent auditory tasks may have been more cognitively demanding, and AC effects may have emerged such that the high AC group would outperform the low AC group. For instance, multiple visual tasks could be presented in separate windows within the interface (see Norcio & Stanley, 1989), and the high AC group may be able to allocate attention more efficiently in a dual- or multitask paradigm.

Response Bias – β (Beta)

In both Experiments 2 and 3, the high AC group consistently demonstrated a higher mean β than the low AC group. This effect occurred regardless of whether participants were given a choice of LOA, as in Experiment 2, or fixed in a particular LOA, as in Experiment 3. Thus, high AC participants were not able to regulate their conservative bias even when allowed to select a preferred LOA. This effect was also especially robust in the auditory modality, wherein the high AC group had a higher β than the low AC group, and the auditory β of the high AC group was higher than their visual β.

Incorporation of the machine decision-aid, however, did regulate β for those with high AC. Thus, the high AC participants used the automation to moderate their response bias; these individuals may have been advantageous in their ability to integrate the stimuli with their own responses as well as those of the machine, as attending to both their own decisions and those of the machine would demand more attentional manipulation.

Modality effects for β emerged. In Experiment 2, the visual β and auditory β were equivalent for the low AC group. However, in Experiment 3, the visual β was significantly higher than the auditory β for the low AC group. Thus, adaptive automation enabled low AC operators to regulate their response bias across modalities. This indicates that the high AC group may have benefited when given a choice of LOA, however, when this choice was removed, as in Experiment 3, modality differences emerged.

The high AC group demonstrated the same modality effects for β in conditions of flexible LOA and fixed LOA. In both Experiments 2 and 3, the auditory β was higher than the visual β. LOA choice did not moderate the response bias of the high AC group.

The medium AC group exhibited the same modality trend as the high AC group, however, the difference in β was less pronounced. The medium AC group was thus less sensitive to modality effects in terms of reporting targets and nons.

Across all tasks, the low AC group exhibited the most consistently neutral response scheme in Experiment 2, where they had a choice of LOAs. When this choice was removed, the low AC group became more biased in both directions (see Figure 74). The high AC group appeared to be generally more neutral when given an LOA choice, with the exception of the auditory modality wherein the bias was especially conservative. The medium AC group was somewhat conservative, but remained more neutral than the high AC group.

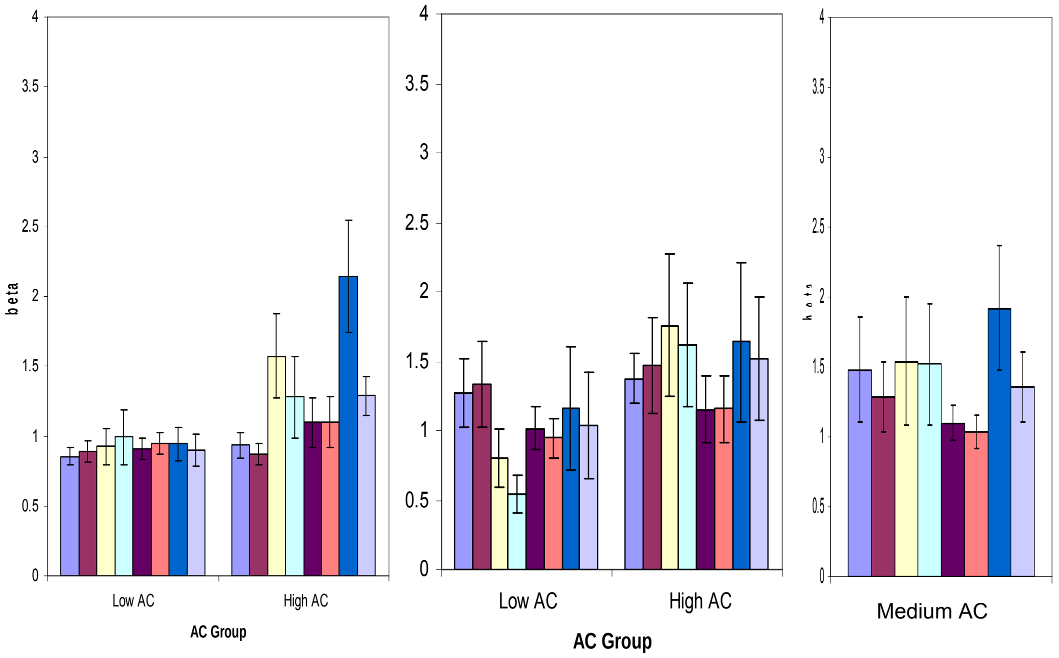


Figure 74. Mean operator-alone and machine-aided β for each of the four tasks across experiments. Experiment 2 is on the left, Experiment 3 is in the middle, and Experiment 4 is on the right.

It may be speculated the LOA choice may benefit the low AC group more than the medium and high AC groups. Low AC participants were efficient in selecting LOAs which enabled them to perceive targets and nons equally well. Their LOA selection strategies enabled them to adapt to the nature of the task, and their attentional needs may be accommodated by a flexible LOA system.

Subjective Workload Ratings

All three AC groups indicated that the combined task generated higher overall workload than the single task conditions. The effects of task load are therefore noted by the operator regardless of attentional control. The medium AC group demonstrated different trends in the subscales than those from the low and high AC groups. In Experiment 2, both the low and high AC groups consistently rated the auditory-only task as less demanding as the visual-only task in all subscales. Experiment 4 showed that the medium AC group gave the auditory-only task higher workload ratings than the visual-only task in terms of temporal demand, effort, and frustration. The medium AC group may therefore not be as subjectively affected by modality differences as the extreme AC groups.

CHAPTER 9: CONCLUSION

Automation has a rich history in aiding the human operator by executing tasks that are difficult and/or undesirable (Lee & See, 2004; Wickens, 1992). While its initial incarnations yielded fixed LOAs to allocate tasks to either the human or machine, adaptive task allocation was advantageous in its ability to reflect changing demands made upon the system and the operator during various phases of the endeavor at hand (Hancock & Chignell, 1989). Parasuraman, Mouloua, and Molloy (1996) reported that observers using nonadaptive automation detected fewer system failures than those using adaptive automation. However, adaptive automation may not always be the ideal arrangement, as the operator may not always be the best judge of the most suitable LOA for the current situation (Morrison & Rouse, 1986), notably when s/he is taxed with many additional task demands that could compete with efficiently managing the choice of LOAs (Sarter & Woods, 1994b; Wiener, 1989). Thus, preferences and needs for various LOAs can at times be incongruous with each other. Further, the operator characteristics which dictate these preferences and needs were yet to be assessed, however, the ability to efficiently allocate attention was hypothesized to be a strong contributor.

The results of Experiments 2 and 4 indicate that an intermediary amount of machine control such as LOA 2 is generally highly preferred. This may be because it gives the operator default control, which encourages task engagement, while also providing confirmation to each of the operators' answers from a machine with 90% reliability. The operators may thus consider 90% reliability to be adequate for stimulus identification suggestions. LOA 4 may have been generally popular in situations where operators approved of the machine's consistent 90% reliable, such as in difficult tasks where the operator preferred to delegate default responsibility to the automation while assuming the less demanding supervisory role. Furthermore, workload assessments from Experiment 3 indicated that LOA 4 generated less workload than LOA 2. Finally, the high AC participants tended to prefer LOA 1, the lowest LOA, more so than the low and medium AC groups.

When preferences were collapsed across tasks, the low AC group was generally more inclined to use LOA 4, the medium AC group was inclined to choose LOA 2, and the high AC group was most inclined to use LOA 1 compared to the other AC groups. A trend was thus observed for specific LOAs to compliment the operators' unique abilities to regulate selective and divided attention as required for various task demands. Thus while AC level may be generally inversely proportional to preferences for automation, there may be other individual difference variables involved. Future research could replicate the methodology of the current experiments, but sample from low, medium, and high desirability of control groups. Desirability of control was positively correlated with AC, and may be one of the factors which encouraged the various AC groups to use certain LOAs. Because d' was not impacted by AC in Experiment 2, it is possible that desirability of control may instead be more influential in LOA preferences and could be assessed in future investigations.

Despite being allowed to choose any LOA, operators performed better in the situations in which there was no choice. This may have been due to the extra task demand of making LOA choices and switching LOAs in Experiment 2. Simply attending to the idea of switching LOAs may have diverted attention away from the target detection tasks. Having an LOA choice may not always be optimal; it is sometimes more beneficial to let the system invoke automation, especially when the operator is too preoccupied to do so (Sarter & Woods, 1994b; Wiener, 1989). Additionally, the operator's ability to assess the need for automation may not always be ideal (Morrison & Rouse, 1986). Thus, in deciding whether to give the operator control over LOAs in a complex system, designers should bear in mind that the amount of control s/he wants may not necessarily reflect that which s/he needs in the present situation. Another important objective is to minimize demands associated with control and supervision of flexible automation (Woods, 1996); this can be fostered by taking care to note the usability of adaptive automation, as well as proper training in the use of various LOAs.

Response bias was strongly impacted by AC. Although the mechanisms driving this behavior are unclear, future research can assess the variability of response bias among operators. Target detection tasks which invoke connotations of threat can be assigned to operators with various levels of AC and TA. Various payoff matrices can be manipulated to influence response bias.

Machine aid, however, can be helpful in regulating response biases. The machine aid in the current series of experiments was used by participants and also effective in approximating response neutrality. Performance can also be improved by a machine aid, provided that it is reasonably reliable and usable. Finally, subjective mental workload can be reduced by using a machine aid; these results are concurrent with the propositions set forth by Endlsey (1999).

The current series of experiments has served as a springboard for future research regarding the impact of individual differences in operators and the consequent implications for automation invocation. While attentional behavior has demonstrated its effects upon operator-automation interaction, additional individual difference variables may have important implications for current automation technology as well as future automation technology which can adapt to a variety of human characteristics to maximize operator performance and safety.

APPENDIX A: ATTENTIONAL CONTROL SCALE

This questionnaire contains 20 statements. Read each statement carefully and decide how well it describes you. For each statement, respond by circling the response that best represents your opinion using the following choices:

Almost Never Sometimes Often Always

1. It's very hard for me to concentrate on a difficult task when there are noises around.

Almost Never Sometimes Often Always

2. When I need to concentrate and solve a problem, I have trouble focusing my attention.

Almost Never Sometimes Often Always

3. When I am working hard on something, I still get distracted by events around me.

Almost Never Sometimes Often Always

4. My concentration is good even if there is music in the room around me.

Almost Never Sometimes Often Always

5. When concentrating, I can focus my attention so that I become unaware of what's going on in the room around me.

Almost Never Sometimes Often Always

6. When I am reading or studying, I am easily distracted if there are people talking in the same room.

Almost Never Sometimes Often Always

7. When trying to focus my attention on something, I have difficulty blocking out distracting thoughts.

Almost Never Sometimes Often Always

8. I have a hard time concentrating when I’m excited about something.

Almost Never Sometimes Often Always

9. When concentrating I ignore feelings of hunger or thirst.

Almost Never Sometimes Often Always

10. I can quickly switch from one task to another.

Almost Never Sometimes Often Always

11. It takes me a while to get really involved in a new task.

Almost Never Sometimes Often Always

12. It is difficult for me to coordinate my attention between the listening and writing required when taking notes during lectures.

Almost Never Sometimes Often Always

13. I can become interested in a new topic very quickly when I need to.

Almost Never Sometimes Often Always

14. It is easy for me to read or write while I’m also talking on the phone.

Almost Never Sometimes Often Always

15. I have trouble carrying on two conversations at once.

Almost Never Sometimes Often Always

16. I have a hard time coming up with new ideas quickly.

Almost Never Sometimes Often Always

17. After being interrupted or distracted, I can easily shift my attention back to what I was doing before.

Almost Never Sometimes Often Always

18. When a distracting thought comes to mind, it is easy for me to shift my attention away from it.

Almost Never Sometimes Often Always

19. It is easy for me to alternate between two different tasks.

Almost Never Sometimes Often Always

20. It is hard for me to break away from one way of thinking about something and look at it from another point of view.

Almost Never Sometimes Often Always

APPENDIX B: DESIRABILITY OF CONTROL SCALE

Instructions: Below you will find a series of statements. Please read each statement carefully and respond to it by expressing the extent to which you believe the statement applies to you. For all items, a response from 1 to 7 is required. Use the number that best reflects your belief when the scale is defined as follows. Write your answers on a separate sheet of paper and click on the "score" button upon completion.

1 = The statement does not apply to me at all

2 = The statement usually does not apply to me

3 = Most often, the statement does no apply

4 = I am unsure about whether or not the statement applies to me, or it applies to me about half the time.

5 = The statement applies more often than not

6 = The statement usually applies to me

7 = The statement always applies to me

____ 1. I prefer a job where I have a lot of control over what I do and when I do it.

____ 2. I enjoy political participation because I want to have as much of a say in running government as possible.

____ 3. I try to avoid situations where someone else tells me what to do.

____ 4. I would prefer to be a leader than a follower.

____ 5. I enjoy being able to influence the actions of others.

____ 6. I am careful to check everything on an automobile before I leave for a long trip.

____ 7. Others usually know what is best for me.

____ 8. I enjoy making my own decisions.

____ 9. I enjoy having control over my own destiny.

____ 10. I would rather someone else take over the leadership role when I am involved in a group project.

____ 11. I consider myself to be generally more capable of handling situations than others are.

____ 12. I would rather run my own business and make my own mistakes than listen to someone else's orders.

____ 13. I like to get a good idea of what a job is all about before I begin.

____ 14. When I see a problem, I prefer to do something about it rather than sit by and let it continue.

____ 15. When it comes to orders, I would rather give them than receive them.

____ 16. I wish I could push many of life's daily decisions off on someone else.

____ 17. When driving, I try to avoid putting myself in a situation where I could be hurt by another person's mistake.

____ 18. I prefer to avoid situations where someone else has to tell me what it is I should be doing.

____ 19. There are many situations in which I would prefer only one choice rather than having to make a decision.

____ 20. I like to wait and see if someone else is going to solve a problem so that I don't have to be bothered by it.

APPENDIX C: STATE-TRAIT ANXIETY INVENTORY

ID: ____________________

DATE: ____________________

Self-Evaluation Questionnaire

DIRECTIONS: A number of statements which people have used to Describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you *generally* feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

	ALMOST NEVER	SOMETIMES	OFTEN	ALMOST ALWAYS
1. I feel pleasant	1	2	3	4
2. I feel nervous and restless	1	2	3	4
3. I feel satisfied with myself	1	2	3	4
4. I wish I could be as happy as others seem to be	1	2	3	4
5. I feel like a failure	1	2	3	4
6. I feel rested	1	2	3	4
7. I am "calm, cool, and collected"	1	2	3	4
8. I feel that difficulties are piling up so that I cannot overcome them	1	2	3	4
9. I worry too much over something that really doesn't matter	1	2	3	4
10. I am happy	1	2	3	4
11. I have disturbing thoughts	1	2	3	4

	ALMOST NEVER	SOMETIMES	OFTEN	ALMOST ALWAYS
12. I lack self-confidence …………………………………………………………..............	1	2	3	4
13. I feel secure ……………………………………………………………………............	1	2	3	4
14. I make decisions easily ……………………………………………………………	1	2	3	4
15. I feel inadequate ……………………………………………………………………...	1	2	3	4
16. I am content ……………………………………………………………………............	1	2	3	4
17. Some unimportant thought runs through my mind and bothers me ……………..	1	2	3	4
18. I take disappointments so keenly that I can't put them out of my mind ………….	1	2	3	4
19. I am a steady person ………………………………………………………………	1	2	3	4
20. I get in a state of tension or turmoil as I think over my recent concerns and interests ……………………………………………………………………….……...	1	2	3	4

APPENDIX D: THE AUTOMATION LEVEL PREFERENCE INDEX

Please read the scenario below and indicate your answer by marking the percentage on the scale which corresponds to your answer.

Scenario:
You are an aircraft pilot responsible for detecting threatening tanks on the ground using a radar system. Both hostile threatening tanks and friendly tanks are on the ground at the same time. You must be very careful to correctly identify which tanks are threatening tanks and which tanks are friendly tanks because they look very similar.

Your aircraft has a computer system which can automatically detect the threatening tanks using artificial intelligence. This computer correctly identifies enemy tanks about 90% of the time. However, it incorrectly identifies friendly tanks as threatening tanks about 10% of the time. Remember that a human operator can also make mistakes in identifying the tanks.

Please rate the degree of computer assistance you would like to use given the scenario you have just read.
Do this by marking the vertical line below at the point that best represents what percentage of the task the computer should do for you. You can mark anywhere along the scale, ranging **anywhere between 0% to 100%.**

100%

I want the computer to make all the decisions and actions **without any input from me**. It is okay if the computer is correct 90% of the time and wrong 10% of the time.

I want the computer to detect all the threatening tanks. I should supervise the computer's decision and be able to **override its decision if I wish to**, since the computer is only correct

50%

I want to share the responsibility **equally** with the computer. **The computer will detect threatening tanks on the right side** of the radar screen, and I can override the computer's decisions if I want. **I will look for threatening tank tanks on the left side** of the radar

I want to detect all of the threatening tanks, but I would like the computer to give me suggestions after each of my decisions and tell me if it thinks it saw a threatening tank, even

I want to make all the decisions and actions **without any input from the computer**, since the computer is correct only 90% of the time and wrong 10% of the time.

0%

APPENDIX E: INFORMED CONSENT (EXPERIMENT 1)

The University of Central Florida and the UCF Department of Psychology support the protection of human subjects participating in research. We are presenting the following information so that you can decide whether you wish to participate in this study.

The proposed research will investigate the role of human operator individual differences in the configuration of adaptive automation systems. The experimenters believe that certain traits of human operators can indicate their needs for certain levels of task automation. Automated systems are used today to complete tasks for humans, and while automation is helpful to human users, not all human users respond to automation in the same way. Some users prefer more or less control over the automation, and it is possible that this preference is influenced by individual differences in human users. The purpose of this study is to investigate the influence of individual differences in human users on preferences and needs of automated systems. For this research study, you will be asked to complete questionnaires regarding individual differences. These individual differences include your general desire for control in everyday settings, attentional control, and trait anxiety. You will be asked to answer questions about yourself. We are not interested in any particular person's specific responses. You are not required to answer or participate in any part of the research. Completion of these questionnaires is estimated to involve approximately forty minutes. You will be allowed to take breaks as needed. Furthermore, all of the data collected in this study will be kept completely confidential and throughout the study, you will be identified by a subject number only. No names will be used in any setting, including any reporting or publication that may result from the data gathered. All of the data collected in this study will be reported in aggregate form only (in other words, participants will be referred to only as numbers, not names). This subject number will not be linked to your name in any way.

The benefit to you is extra credit and added knowledge about participation in psychological research. You will be compensated for your participation in this research through experimental credit for courses. Your participation is strictly voluntary and you may withdraw at any time without negative consequence.

You may be offered the opportunity to participate in further related studies and receive payment for your participation. We will be conducting a future study related to this experiment within the next two months. This research will investigate the role of human traits in the use of computerized adaptive automation systems. You are NOT required to participate in this study. However, if you would like to participate, you may complete the form which will be provided to you at the end of the study. For this future study, you will be paid at the rate of $7.50 per hour. It is estimated that this future study will take approximately 2 and a half hours to complete, and so you will receive $20. Your participation is strictly voluntary and you may withdraw at any time without negative consequence. There are no anticipated risks or discomforts to you as a subject in this future study. Your trait data from this study will be used to analyze the results of your responses from study 2. You will still be granted anonymity such that your name and contact information will never appear linked to your performance and trait analyses in the second and third studies. Your name will be linked only by a coding sheet that will be kept separately from your data and responses. After the study is complete, the coding sheet will be physically destroyed, and any electronic files (such as an electronic version of the coding sheet) that may have linked your name to your data will be erased. Your contact information will never be accessible to anybody except the experimenter. All research information gathered will be reported in aggregate form only (meaning participants will be referred to as numbers only) and your name will not be used in the reporting or publications that may result from the data gathered.

If you wish to see the results of this study, you may request a write-up of them from the investigators listed below. Additionally, you may contact the investigator with questions about this research.

Primary Investigators:

Jennifer Thropp
Dept. of Psychology
Univ. of Central Florida
(386) 451-6948
j_thropp@hotmail.com

Tal Oron-Gilad
Dept. of Psychology
Univ. of Central Florida
(407) 823-0923
torongil@mail.ucf.edu

The University of Central Florida requires that the following statement appear on all informed consent forms.

If you believe you have been injured during participation in this research project, you may file a claim with UCF Environmental Health & Safety, Risk and Insurance Office, P.O. Box 163500, Orlando, FL 32816-3500, (407) 823-6300. The University of Central Florida is an agency of the State of Florida for purposes of sovereign immunity and the university's and the state's liability for personal injury or property damage is extremely limited under Florida law. Accordingly, the university's and the state's ability to compensate you for any personal injury or property damage suffered during this research project is very limited.

Information regarding you rights as a research volunteer may be obtained from:

Barbara Ward
Institutional Review Board (IRB)
University of Central Florida (UCF)
12443 Research Parkway, Suite 302
Orlando, Florida 32826-3252
Telephone: (407) 823-2901

I have read the procedure described above. I voluntarily agree to participate in the procedure, and I have received a copy of this description. I am 18 years of age or older AND I AM CAPABLE OF PROVIDING CONSENT TO PARTICIPATE. I agree to participate in this study.

_______________________________________ ________________________

Signature of Participant Date

APPENDIX F: AGREEMENT TO BE CONTACTED FOR PARTICIPATION IN FUTURE STUDY

Thank you for your participation in this study. We will be conducting a future study related to this experiment within the next two months. This research will investigate the role of human traits in the use of computerized adaptive automation systems. You are not required to participate in this study. However, if you would like to participate, please read the following information and complete the contact information below.

For this future study, you will be paid at the rate of $7.50 per hour. It is estimated that it will take approximately two and a half hours to complete. There are no anticipated risks or discomforts to you as a subject in this study. Your trait data from this study will be used to analyze the results of your responses from the second study. You will still be granted anonymity such that your name and contact information will never appear linked to your performance and trait analyses. Your name will be linked only by a coding sheet that will be kept separately from your data and responses. After the study is complete, the coding sheet will be destroyed. Your contact information will never be accessible to anybody except the experimenter.

<u>Second Study Description</u>: You will be asked to complete a computer-based task regarding your feelings for an automated computer system. You will be presented with a computerized automation interface and asked to indicate which of the two target detection tasks you wish to automate (in other words, you will tell the computer if you would like to do the target detection tasks yourself or if you would like the computer to do them for you). You must have normal or corrected vision and you must have normal hearing in both ears (no known hearing impairments).

If you wish to participate in these future studies, please read the agreement and complete the form below:

I have read the procedure described above. I voluntarily agree to be contacted by an experimenter to participate in these future procedures.

Printed Name:____

Best way to reach you (e-mail address, phone number, etc.)_______________________

____

____ ____________
Signature of Participant Date

APPENDIX G: SHORT FORM OF THE NASA-TLX

RATING SHEET

INSTRUCTIONS: On each scale, place a mark that represents the magnitude of that factor in the task you just performed.

How much mental activity was required (thinking, deciding, calculating, remembering, looking, searching)? Was the task easy or demanding?

I--I
LOW HIGH

MENTAL DEMAND

How much physical activity was required (pushing, pulling, turning, controlling, activating)? Was the task easy or demanding, slow or brisk, slack or strenuous?

I--I
LOW HIGH

PHYSICAL DEMAND

How much time pressure did you feel due to the rate or pace at which the task or parts of the task occurred? Was the pace slow and leisurely or rapid and frantic?

I--I
LOW HIGH

TEMPORAL DEMAND

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance?

I--I
LOW HIGH

PERFORMANCE

How hard did you have to work (mentally and/ or physically) to accomplish your level of performance?

I--I
LOW HIGH

EFFORT

How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

I--I
LOW HIGH

FRUSTRATION

APPENDIX H: EXPERIMENT 2 TASK INSTRUCTIONS

You are the operator of a machine and you are responsible for identifying visual and auditory targets. You can use this machine to help you to identify targets if you wish. You can delegate responsibilities to the machine depending on your needs and/or preferences.

You will **see vertical lines** on the screen and **hear tones** in your headphones. Lines will appear on the left and right sides of the screen, and the tones will be presented in the left and right earphones separately. Therefore, there will be four types of stimuli, but they will never occur at the same time. They will occur in random orders.

Your task is to identify each stimulus as a “Target’ or ‘Non' (non-target) as follows:

Visual task stimuli:

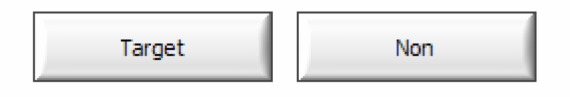
- **Target line**: presented on the screen for the **shorter** period of time (125 ms)
- **Non line**: presented on the screen for the **longer** period of time (250 ms)

Auditory task stimuli:

- **Target tone**: presented in the earphone for the **shorter** period of time (200 ms)
- **Non tone**: presented in the earphone for the **longer** period of time (250 ms)

The term ‘non’ is used instead of ‘nontarget’.

You will identify each stimulus type after it is presented by **using a mouse to click the buttons** labeled ‘Target’ or ‘Non’, depending on your answer.



You will first view a demonstration of how the target and non lines look.

Then you will view a demonstration of how the target and non tones sound.

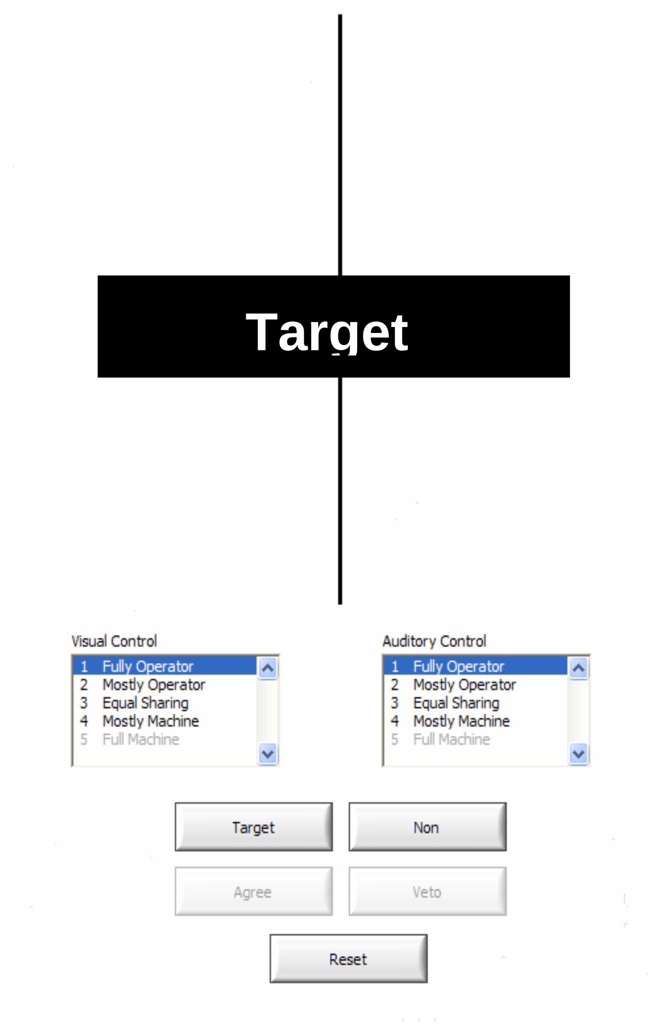
OPTIONAL MACHINE DECISION AID

You have the choice of using the machine's decision aid to help you. There are different levels of machine assistance which we will review later.

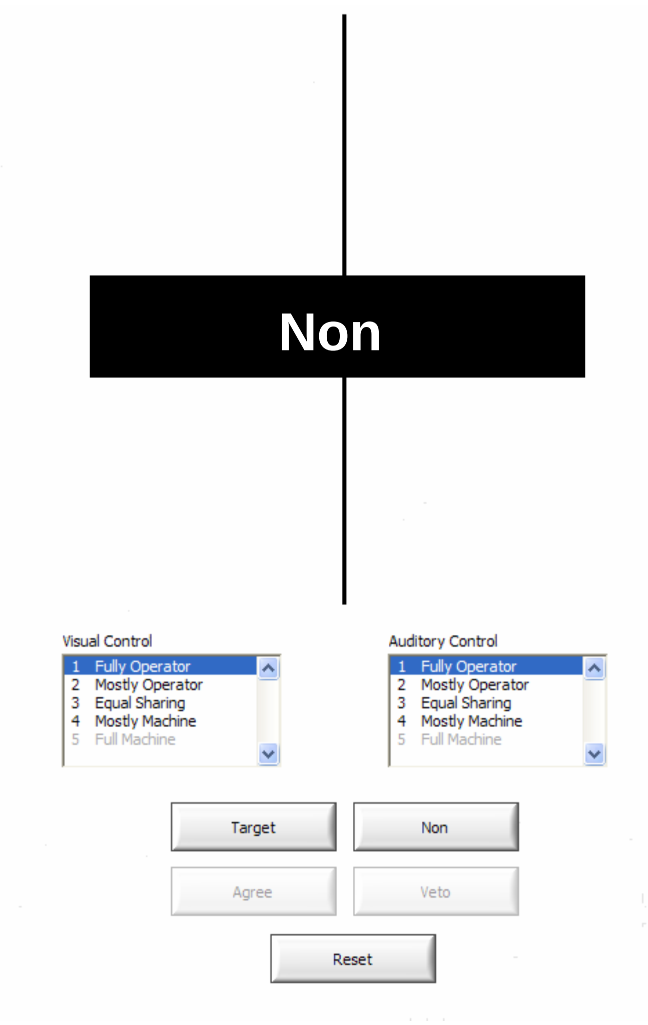
If you use it, the machine will provide a **message on the screen** that reads 'Target' after a target line appears and 'Non' after a non line appears in the visual task.

The machine will also provide a **voice in the headphones** that says 'Target' after a target tone and 'Non' after a non tone in the auditory task.

However, the machine is **90% reliable** in its ability to make correct identifications. This means it identifies a stimulus correctly 90% of the time and incorrectly 10% of the time. It will make some mistakes.



Example of machine response to a target line in the visual task



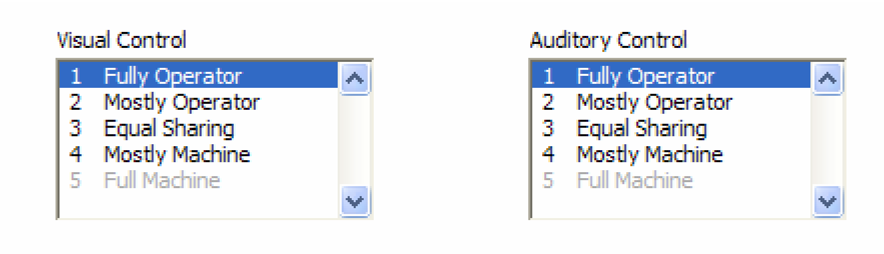
Example of machine response to a non line in the visual task

Control Modes for Machine Assistance

Visual Control Mode: The level of machine assistance for identifying the **lines**.

Auditory Control Mode: The level of machine assistance for identifying the **tones**.

The level of machine assistance which is currently engaged is highlighted on the Auditory Control and the Visual Control menus on the screen. In the example below, both the Visual Control and Auditory Control modes are set at the '1 Fully Operator' level.



The sign on the desk summarizes the **five different levels** of machine assistance available to you for both the Visual Control and Auditory Control modes. You will practice the task at each of these levels to become familiar with them and determine which one you prefer to use during the task.

LEVELS OF MACHINE ASSISTANCE

1: Full Operator:
You must identify all stimuli without any machine suggestions

2: Mostly Operator:
You must first identify the stimulus, and then the machine provides its suggestion, which you can use instead

3: Equal Sharing:
Left screen/ear: You must first identify the stimulus, then the machine provides its answer, which you use can use instead
Right screen/ear: The machine first identifies the stimulus, which you can veto

4: Mostly Machine:
The machine first identifies the stimulus, which you can veto

5: Full Machine:
The machine will identify all stimuli without any input from you

PRACTICE SESSIONS

5 FULL MACHINE CONTROL

The highest level of machine assistance is ‘5 Full Machine’.

5 Full Machine: The machine will identify all stimuli all by itself, so you do not have to do anything except make sure it still functions. You will hear its voiced answers when it identifies a tone, and you will see its message answers when it identifies a line.

You are still responsible for making sure the machine is making its decisions by checking for its answer messages on the screen and voice messages in the earphones. You do not have to pay attention to the lines or tones, but if you think the machine decision aid has stopped working properly (it makes too many mistakes or quits providing answers), click on ‘Reset’ to reactivate it.

The machine is **90% reliable** in its decisions. It will make some mistakes.

Do you have any questions?

You will now view a demonstration of the ‘5 Full Machine’ level in the Visual Control Mode and Auditory Control Mode.

You do not need to practice the ‘5 Full Machine’ level of assistance, because the machine automatically makes all of the identifications and you do not have to do anything.

4 MOSTLY MACHINE CONTROL

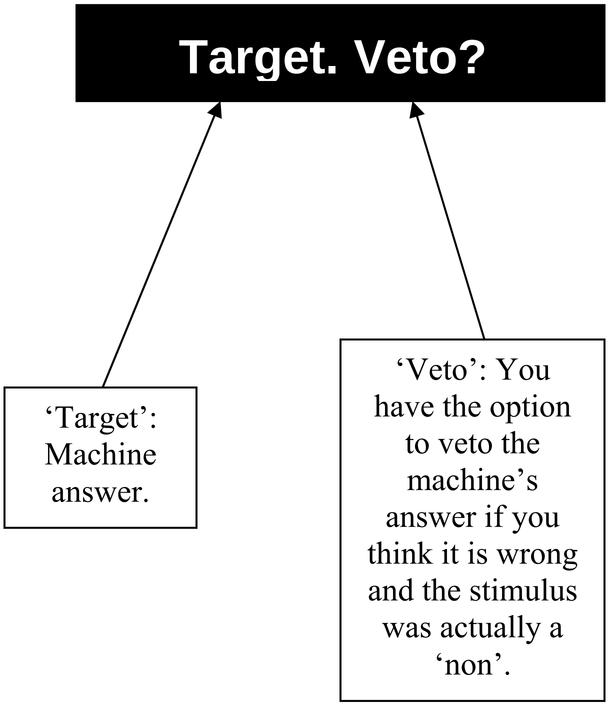
4 Mostly Machine: The machine makes all identification decisions for you, but you can change its answers if you wish.

The machine will identify the stimulus as a ‘target’ or ‘non’ immediately after it is presented.

If you wish, you can check the machine’s answers to make sure they are correct, because the machine can make mistakes.

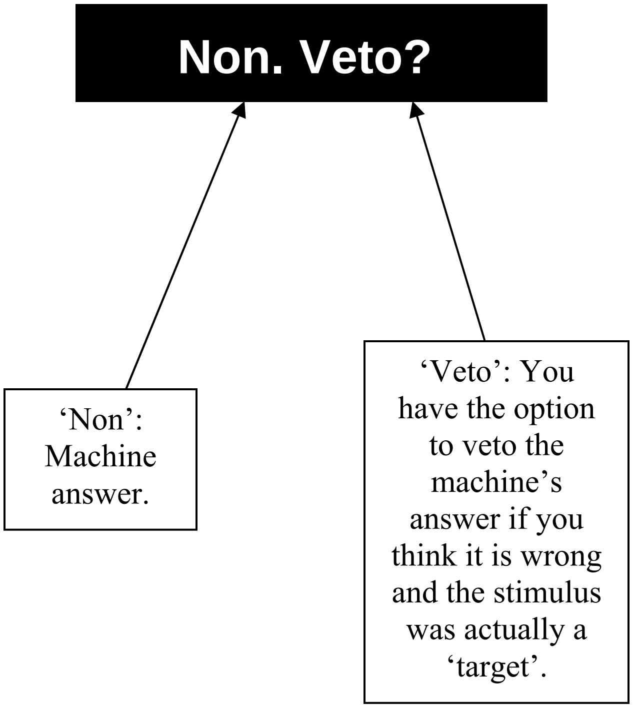
If you think the machine is correct, you can either **click on ‘Agree’ or simply do nothing** and then the machine’s answer will count as your answer. If you think the machine is incorrect, you can click on the **‘Veto’** button to correct its suggestion. If you do not click on “Agree’ or “Veto’, the **machine assumes you agree** with it, and its answer will count towards your score.
You must click on ‘Agree’ or ‘Veto’ within 3 seconds if you choose to do so.

<u>When the stimulus is a line:</u>



‘Target’: Machine answer.

‘Veto’: You have the option to veto the machine’s answer if you think it is wrong and the stimulus was actually a ‘non’.



‘Non’: Machine answer.

‘Veto’: You have the option to veto the machine’s answer if you think it is wrong and the stimulus was actually a ‘target’.

When the stimulus is a tone:

If a target tone is presented, the machine voice in the headphones will say, **"Target. Veto?"** You have the option to veto the machine's answer if you think it is wrong and the stimulus was actually a 'non'.

If a non tone is presented, the machine voice in the headphones will say, **"Non. Veto?"** You have the option to veto the machine's answer if you think it is wrong and the stimulus was actually a 'target'.

Clicking on 'Agree' or 'Veto' after the machine provides its answer is **always optional**. You can either respond to the machine's answers or ignore them, depending on how much you trust the machine's decisions, how much you want to interact with the system, and how much of the task you want to control. You can choose to either respond to the machine's answers or ignore them, but your task is to ensure the highest possible number of correct identifications.

Do you have any questions?

You will now view a demonstration of the '4 Mostly Machine' level in the Visual Control Mode and Auditory Control Mode.

Now you will practice with the identifying the lines and tones as targets or nons at the ‘4 Mostly Machine’ level.

Visual Task Practice settings:

- **Visual Control Mode: 4 Mostly Machine**
- **Auditory Control Mode: 5 Full Machine**

Auditory Task Practice settings:

- **Visual Control Mode: 5 Full Machine**
- **Auditory Control Mode: 4 Full Machine**

3 EQUAL SHARING

3 Equal Sharing: The identification responsibility is divided between you and the machine.

If the stimulus is a line on the **left side** of the screen or a tone in the **left earphone**:
You first identify the stimulus as a 'target' or 'non' within 3 seconds, then the machine will provide its suggestion.

If you like your original answer and think you are correct, then ignore the machine suggestion and wait for the next stimulus.

However, you can **change your answer** based on the machine's suggestion, **if you wish**.
.
If you think the machine is correct and want to use the machine's suggestion as your answer, click on the **'Agree' button**. If you think the machine is incorrect, you can click on the **'Veto' button** to veto the machine's identification.
You must click on 'Agree' or 'Veto' within 3 seconds if you choose to do so.

If the stimulus is a line on the **right side** of the screen or a tone in the **right earphone**:
The machine will first identify the stimulus as a 'target' or 'non' immediately after it is presented.

You do not have to make any identifications, but **if you wish**, you can check the machine's answers to make sure they are correct, because the machine can make mistakes.

If you think the machine is correct and want the machine's answer to count toward your score, click on the **'Agree' button or simply do nothing**. If you think the machine is incorrect, you can click on the **'Veto' button** to correct the machine's identification. If you do not click on "Agree' or "Veto', the **machine assumes you agree** with it, and its answer will count towards your score.
You must click on 'Agree' or 'Veto' within 3 seconds if you choose to do so.

Clicking on **'Agree'** or **'Veto'** after the machine provides its suggestion is **always optional**. If you like your original answer, then you can ignore the machine's answers and you do not need to click 'Agree' or 'Veto'.

In summary, if the stimulus is on the left, you will identify it first. If the stimulus is on the right, the machine will identify it first.

Do you have any questions?

You will now view a demonstration of the ‘3 Equal Sharing’ level in the Visual Control Mode and the Auditory Control Mode.

Now you will practice with the identifying the lines and tones as targets or nons at the ‘3 Equal Sharing’ level.

Visual Task Practice settings:

- **Visual Control Mode: 3 Equal Sharing**
- **Auditory Control Mode: 5 Full Machine**

Auditory Task Practice settings:

- **Visual Control Mode: 5 Full Machine**
- **Auditory Control Mode: 3 Equal Sharing**

2 MOSTLY OPERATOR CONTROL

2 Mostly Operator: After the stimulus appears, you must first identify the stimulus as a ‘target’ or ‘non’ within 3 seconds, and then the machine will provide its suggestion.

If you like your original answer and think you are correct, then ignore the machine suggestion and wait for the next stimulus.

However, you can **change your answer** based on the machine's suggestion, **if you wish**.

If you think the machine is correct and want the machine’s suggestion to be used as your answer, click on the **‘Agree’ button**. If you think the machine is incorrect, you can click on the **‘Veto’ button** to correct the machine’s identification.

You must click on ‘Agree’ or ‘Veto’ within 3 seconds if you choose to do so.

Clicking on **‘Agree’** or **‘Veto’** after the machine provides its suggestion is **always optional**. If you like your original answer, then you can ignore the machine’s suggestions and you do not need to click ‘Agree’ or ‘Veto’.

Remember that the machine is **90% reliable** in its decisions. It will make some mistakes.

Do you have any questions?

You will now view a demonstration of the ‘2 Mostly Operator’ level in the Visual Control Mode and Auditory Control Mode.

Now you will practice with the identifying the lines and tones as targets or nons at the '2 Mostly Operator' level.

Visual Task Practice settings:

- **Visual Control Mode: 2 Mostly Operator**
- **Auditory Control Mode: 5 Full Machine**

Auditory Task Practice settings:

- **Visual Control Mode: 5 Full Machine**
- **Auditory Control Mode: 2 Mostly Operator**

1 FULL OPERATOR CONTROL

1 Full Operator. You must identify all stimuli within 3 seconds without any suggestions from the machine.

After you identify a stimulus, the next stimulus will be presented.

Do you have any questions?

You will now view a demonstration of the ‘1 Full Operator’ level in the Visual Control Mode and Auditory Control Mode.

Now you will practice with the identifying the lines and tones as targets or nons at the '1 Full Operator' level.

Visual Task Practice settings:

- **Visual Control Mode: 1 Full Operator**
- **Auditory Control Mode: 5 Full Machine**

Auditory Task Practice settings:

- **Visual Control Mode: 5 Full Machine**
- **Auditory Control Mode: 1 Full Operator**

THREE TEST SESSIONS

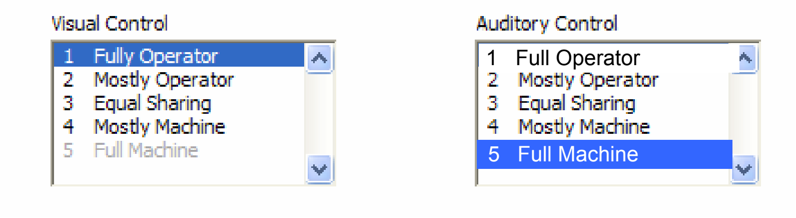
TEST SESSION: IDENTIFYING ONLY LINES

Now you will begin a test session, in which you are responsible for identifying the **lines only** as targets or nons.

During this entire session, you can change the level of machine assistance **as often as you wish** for the Visual Control Mode.

You can change a level of machine assistance by clicking on your choice from the menu list. **You may choose only among levels 1 Full Operator, 2 Mostly Operator, 3 Equal Sharing, and 4 Mostly Machine for the Visual Control Mode.**

You cannot select '5 Full Machine' for the Visual Control Mode because you must control at least some portion of the line identification task. You will see that the '5 Full Machine' choice on the menu for Visual Control Mode is disabled.



<u>Test settings:</u>

- **Visual Control Mode: <u>Your choice</u> of levels:**
  1 Full Operator,
  2 Mostly Operator,
  3 Equal Sharing, and
  4 Mostly Machine
- Auditory Control Mode: Fixed at 5 Full Machine (machine will identify tones all by itself, so you can ignore the tones)

If you think the machine decision aid has failed in its tone identifications, click on 'Reset' to reactivate it.

Remember that the machine is **90% reliable** in its decisions. It will make some mistakes.

Please use the level of machine assistance for the Visual Control Mode which you feel will result in the **best performance**. This will depend on your own ability to correctly identify target and non lines, as well as your preference for the machine assistance. You can change the level as often as you wish during the entire session until you find a level at which you are comfortable.

Your final answer for each stimulus will be used to compute your score, **whether you use machine assistance or not**. We are interested in knowing which level of assistance helps you **perform the best and is your preference**.

The participant with the best overall performance will receive a **$100 cash reward**.

Do you have any questions?

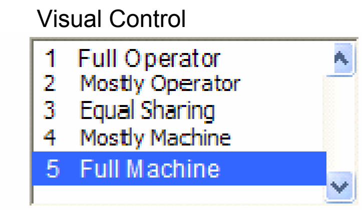
TEST SESSION: IDENTIFYING ONLY TONES

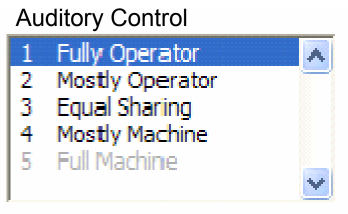
Now you will begin a test session, in which you are responsible for identifying the **tones only** as targets or nons.

During this session, you can change the level of machine assistance **as often as you wish** for the Auditory Control Mode.

You can change a level of machine assistance by clicking on your choice from the menu list. **You may choose only among levels 1 Full Operator, 2 Mostly Operator, 3 Equal Sharing, and 4 Mostly Machine for the Auditory Control Mode.**

You cannot select '5 Full Machine' for the Auditory Control Mode because you must control at least some portion of the tone identification task. You will see that the '5 Full Machine' choice on the menu for Auditory Control Mode is disabled.





<u>Test settings:</u>

- Visual Control Mode: 5: Fixed Full Machine (the machine will identify all lines, so you can ignore the lines)
- **Auditory Control Mode: <u>Your choice</u> of levels:**
  1 Full Operator,
  2 Mostly Operator,
  3 Equal Sharing, and
  4 Mostly Machine

If you think the machine decision aid has failed in its line identifications, click on 'Reset' to reactivate it.

Remember that the machine is **90% reliable** in its decisions. It will make some mistakes.

Please use the level of machine assistance for the Auditory Control Mode which you feel will result in the **best performance**. This will depend on your own ability to correctly identify target and non tones, as well as your preference for the machine assistance. You can change the level as often as you wish during the entire session until you find a level at which you are comfortable.

Your final answer for each stimulus will be used to compute your score, **whether you use machine assistance or not**. We are interested in knowing which level of assistance helps you **perform the best and is your preference**.

The participant with the best overall performance will receive a **$100 cash reward**.

Do you have any questions?

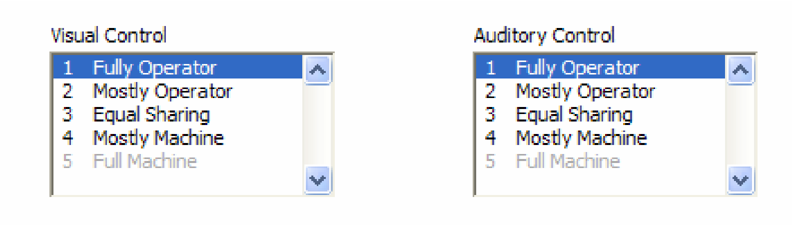
TEST SESSION: IDENTIFYING BOTH LINES AND TONES

Now you will begin the final session, in which you are responsible for identifying the both the **lines and tones** as targets or nons.

During this session, you can change the level of machine assistance **as often as you wish** for both the Visual Control Mode and the Auditory Control Mode.

You can change the level of machine assistance by clicking on your choice from the menu list. **You can choose only among levels 1 Full Operator, 2 Mostly Operator, 3 Equal Sharing, and 4 Mostly Machine for both the Visual Control Mode and the Auditory Control Mode.**

You cannot select '5 Full Machine' for Visual Control Mode or Auditory Control Mode because you must control at least some portion of the line and tone identification tasks. You will see that the ‘5 Full Machine’ choice on the menu for both Visual Control Mode and Auditory Control Mode is disabled.



Test settings:

- **Visual Control Mode: <u>Your choice</u> of levels:**
    1 Full Operator,
    2 Mostly Operator,
    3 Equal Sharing, and
    4 Mostly Machine
- **Auditory Control Mode: <u>Your choice</u> of levels:**
    1 Full Operator,
    2 Mostly Operator,
    3 Equal Sharing, and
    4 Mostly Machine

Remember that the machine is **90% reliable** in its decisions. It will make some mistakes.

Please use the level of machine assistance for the Visual Control Mode and Auditory Control Mode which you feel will result in the **best performance**. This will depend on your own ability to correctly identify target and non lines and tones, as well as your preference for the machine assistance.

Your final answer for each stimulus will be used to compute your score, **whether you use machine assistance or not**. We are interested in knowing which level of assistance helps you **perform the best and is your preference**. You can change the level as often as you wish during the entire session until you find a level at which you are comfortable.

The participant with the best overall performance will receive a **$100 cash reward**.

Do you have any questions?

APPENDIX I: EXPERIMENT 2 INFORMED CONSENT

The University of Central Florida and the UCF Department of Psychology support the protection of human subjects participating in research. We are presenting the following information so that you can decide whether you wish to participate in this study.

The proposed research will investigate the role of human operator individual differences in the configuration of adaptive automation systems. The experimenters believe that certain traits of human operators can indicate their needs for certain levels of task automation. For this research study, you will be asked to complete a computer-based task about your preferences for an automated system. You will be presented with a computerized automation interface on a computer monitor and asked to indicate which of the three target detection tasks you wish to automate. In other words, you will tell the computer which of the two target detection tasks that you would like to either do yourself or have the computer do instead. You will indicate your response by selecting a button on the monitor. The target detection tasks will be visual and auditory in nature. We are not interested in any particular person's specific responses. You are not required to answer or participate in any part of the research. Furthermore, all of the data collected in this study will be kept completely confidential and throughout the study, you will be identified by a subject number only. No names will be used in any setting, including any reporting or publication that may result from the data gathered. All of the data collected in this study will be reported in aggregate form only (in other words, participants will be referred to only as numbers, not names). This subject number will not be linked to your name in any way.

There are no anticipated risks or discomforts to you as a subject in this study. The use of this simulator may be accompanied by a mild risk of symptoms similar to motion sickness, including disorientation, nausea, and dizziness. To minimize the potential of such problems occurring, the room will be kept at an appropriate temperature, short duration tasks and regularly scheduled breaks in the use of the simulator will be provided, a reorientation period will be provided after the use of the simulator is complete, and you may cease use of the simulator if you begin to experience any adverse symptoms. If you experience any of these symptoms, report them to the experimenter immediately, and your participation in the study will be stopped without any penalty to you. If the experimenter observes any adverse symptoms occurring, he or she will immediately halt your use of the equipment.

Completion of this study estimated to involve two and a half hours of participation time.

The benefit to you is added knowledge about participation in psychological research. You will be compensated for your participation in this research through experimental credit for courses and/or payment at the rate of $7.50 per hour and you will receive a payment voucher for $20 total. Your participation is strictly voluntary and you may withdraw at any time without negative consequence.

If you wish to see the results of this study, you may request a write-up of them from the investigators listed below. Additionally, you may contact the investigator with questions about this research.

Primary Investigators:

Jennifer Thropp	Tal Oron-Gilad
Dept. of Psychology	Dept. of Psychology
Univ. of Central Florida	Univ. of Central Florida
(386) 451-6948	(407) 823-0923
j_thropp@hotmail.com	torongil@mail.ucf.edu

The University of Central Florida requires that the following statement appear on all informed consent forms.

If you believe you have been injured during participation in this research project, you may file a claim with UCF Environmental Health & Safety, Risk and Insurance Office, P.O. Box 163500, Orlando, FL 32816-3500, (407) 823-6300. The University of Central Florida is an agency of the State of Florida for purposes of sovereign immunity and the university's and the state's liability for personal injury or property damage is extremely limited under Florida law. Accordingly, the university's and the state's ability to compensate you for any personal injury or property damage suffered during this research project is very limited.

Information regarding you rights as a research volunteer may be obtained from:

Barbara Ward
Institutional Review Board (IRB)
University of Central Florida (UCF)
12443 Research Parkway, Suite 302
Orlando, Florida 32826-3252
Telephone: (407) 823-2901

I have read the procedure described above. I voluntarily agree to participate in the procedure, and I have received a copy of this description. I am 18 years of age or older AND I AM CAPABLE OF PROVIDING CONSENT TO PARTICIPATE. I agree to participate in this study.

____________________________________ ____________
Signature of Participant Date

APPENDIX J: LOA VISUAL AID

LEVELS OF MACHINE ASSISTANCE

1: Full Operator:
You must identify all stimuli without any machine suggestions

2: Mostly Operator:
You must first identify the stimulus, and then the machine provides its suggestion, which you can use instead

3: Equal Sharing:
Left screen/ear: You must first identify the stimulus, then the machine provides its answer, which you use can use instead
Right screen/ear: The machine first identifies the stimulus, which you can veto

4: Mostly Machine:
The machine first identifies the stimulus, which you can veto

5: Full Machine:
The machine will identify all stimuli without any input from you

APPENDIX K: DEBRIEFING

Debriefing Statement

The proposed research has investigated the role of human operator individual differences in the configuration of adaptive automation systems.

Please keep the details of this experiment confidential in order to avoid influencing future participants, thus preserving the results. If you have any questions, please feel free to contact the researchers:

Jennifer E. Thropp, M.S.
Department of Psychology
P.O. Box 161390
Orlando, FL 32816
Telephone: (386) 451-6948
Fax: 407-823-0921
Email: j_thropp@hotmail.com

Or

Tal Oron-Gilad, Ph.D.
Department of Psychology
University of Central Florida
P.O. Box 161390
Orlando, FL 32816
Telephone: (407) 823-0923
Fax: 407-823-0921
Email: torongil@mail.ucf.edu

APPENDIX L: EXPERIMENT 3 INFORMED CONSENT

The University of Central Florida and the UCF Department of Psychology support the protection of human subjects participating in research. We are presenting the following information so that you can decide whether you wish to participate in this study.

The proposed research will investigate the role of human operator individual differences in the configuration of adaptive automation systems. The experimenters believe that certain traits of human operators can indicate their needs for certain levels of task automation. For this research study, you will be asked to complete a computer-based task about your preferences for an automated system. You will be presented with a computerized automation interface on a computer monitor and asked to indicate which of the three target detection tasks you wish to automate. In other words, you will tell the computer which of the two target detection tasks that you would like to either do yourself or have the computer do instead. You will indicate your response by selecting a button on the monitor. The target detection tasks will be visual and auditory in nature. We are not interested in any particular person's specific responses. You are not required to answer or participate in any part of the research. Furthermore, all of the data collected in this study will be kept completely confidential and throughout the study, you will be identified by a subject number only. No names will be used in any setting, including any reporting or publication that may result from the data gathered. All of the data collected in this study will be reported in aggregate form only (in other words, participants will be referred to only as numbers, not names). This subject number will not be linked to your name in any way.

There are no anticipated risks or discomforts to you as a subject in this study. The use of this simulator may be accompanied by a mild risk of symptoms similar to motion sickness, including disorientation, nausea, and dizziness. To minimize the potential of such problems occurring, the room will be kept at an appropriate temperature, short duration tasks and regularly scheduled breaks in the use of the simulator will be provided, a reorientation period will be provided after the use of the simulator is complete, and you may cease use of the simulator if you begin to experience any adverse symptoms. If you experience any of these symptoms, report them to the experimenter immediately, and your participation in the study will be stopped without any penalty to you. If the experimenter observes any adverse symptoms occurring, he or she will immediately halt your use of the equipment.

Completion of this study estimated to involve two and a half hours of participation time.

The benefit to you is added knowledge about participation in psychological research. You will be compensated for your participation in this research through experimental credit for courses and/or payment at the rate of $7.50 per hour and you will receive a payment voucher for $20 total. Your participation is strictly voluntary and you may withdraw at any time without negative consequence.

If you wish to see the results of this study, you may request a write-up of them from the investigators listed below. Additionally, you may contact the investigator with questions about this research.

Primary Investigators:

Jennifer Thropp
Dept. of Psychology
Univ. of Central Florida
(386) 451-6948
j_thropp@hotmail.com

Tal Oron-Gilad
Dept. of Psychology
Univ. of Central Florida
(407) 823-0923
torongil@mail.ucf.edu

The University of Central Florida requires that the following statement appear on all informed consent forms.

If you believe you have been injured during participation in this research project, you may file a claim with UCF Environmental Health & Safety, Risk and Insurance Office, P.O. Box 163500, Orlando, FL 32816-3500, (407) 823-6300. The University of Central Florida is an agency of the State of Florida for purposes of sovereign immunity and the university's and the state's liability for personal injury or property damage is extremely limited under Florida law. Accordingly, the university's and the state's ability to compensate you for any personal injury or property damage suffered during this research project is very limited.

Information regarding you rights as a research volunteer may be obtained from:

Barbara Ward
Institutional Review Board (IRB)
University of Central Florida (UCF)
12443 Research Parkway, Suite 302
Orlando, Florida 32826-3252
Telephone: (407) 823-2901

I have read the procedure described above. I voluntarily agree to participate in the procedure, and I have received a copy of this description. I am 18 years of age or older AND I AM CAPABLE OF PROVIDING CONSENT TO PARTICIPATE. I agree to participate in this study.

Signature of Participant	Date

APPENDIX M: EXPERIMENT 3 TASK INSTRUCTIONS FOR PARTICIPANTS FIXED IN LOA 2

You are the operator of a machine and you are responsible for identifying visual and auditory targets.

You will **see vertical lines** on the screen and **hear tones** in your headphones. Lines will appear on the left and right sides of the screen, and the tones will be presented in the left and right earphones separately. Therefore, there will be four types of stimuli, but they will never occur at the same time. They will occur in random orders.

Your task is to identify each stimulus as a “Target’ or ‘Non' (non-target) as follows:

Visual task stimuli:

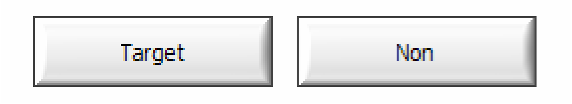
- **Target line**: presented on the screen for the **shorter** period of time (125 ms)
- **Non line**: presented on the screen for the **longer** period of time (250 ms)

Auditory task stimuli:

- **Target tone**: presented in the earphone for the **shorter** period of time (200 ms)
- **Non tone**: presented in the earphone for the **longer** period of time (250 ms)

The term ‘non’ is used instead of ‘nontarget’.

You will identify each stimulus type after it is presented by **using a mouse to click the buttons** labeled ‘Target’ or ‘Non’, depending on your answer.



You will first view a demonstration of how the target and non lines look.

Then you will view a demonstration of how the target and non tones sound.

The machine will identify all the stimuli immediately after they are presented.

MACHINE DECISION AID

The machine's decision aid will assist you in identifying all targets and nons. There are two different levels of machine assistance which you will use during the tasks.

They are:

5 Full Machine: Machine will make all target and non identifications automatically without any input from you

and

2 Mostly Operator: You must first identify each stimulus as a target or non, and then the machine will provide its suggestion, which you can then use as your answer instead

The machine is always **90% reliable** in its ability to make correct identifications. This means it identifies a stimulus correctly 90% of the time and incorrectly 10% of the time. It will make some mistakes.

There are three other levels of assistance which the machine can offer, but you will not be using them.

Now you will learn how to use the two levels which will be available to you, which are "5 Full Machine" and "2 Mostly Operator".

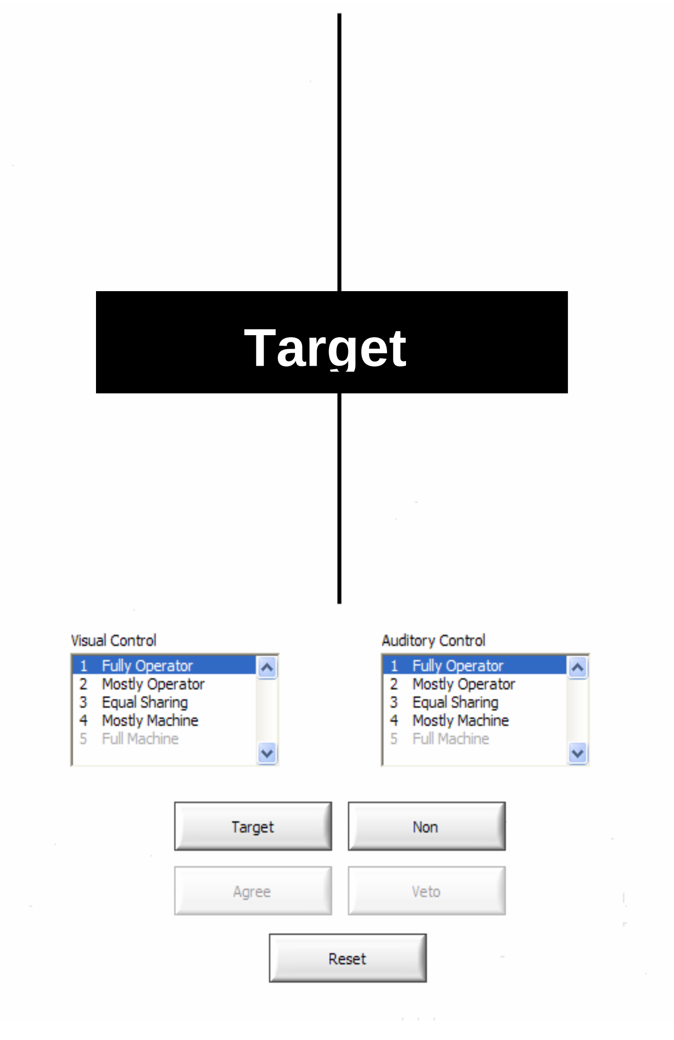
5 FULL MACHINE CONTROL

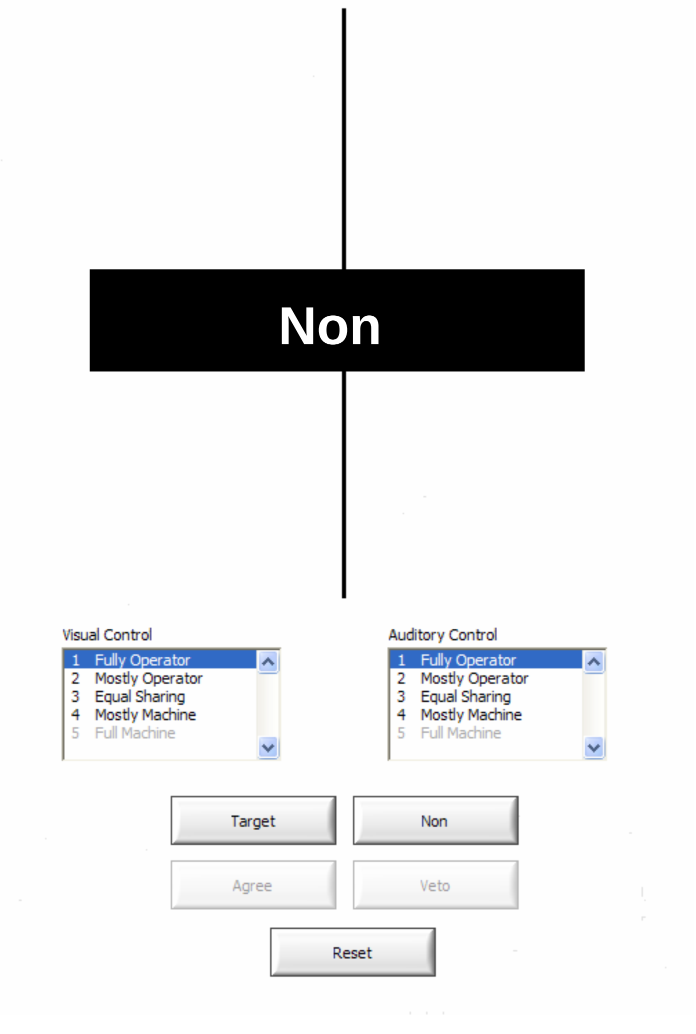
The highest level of machine assistance is ‘5 Full Machine’.

5 Full Machine: The machine makes all identifications automatically, so you do not have to do anything except make sure it still functions. You will hear its voiced answers when it identifies a tone, and you will see its message answers when it identifies a line.

The machine will provide a **message on the screen** that reads ‘Target’ after a target line appears and ‘Non’ after a non line appears in the visual task.

The machine will also provide a **voice in the headphones** that says ‘Target’ after a target tone and ‘Non’ after a non tone in the auditory task.





You are still responsible for making sure the machine is making its decisions by checking for its answer messages on the screen and voice messages in the earphones. You do not have to pay attention to the lines or tones, but if you think the machine decision aid has stopped working properly (it makes too many mistakes or quits providing answers), click on ‘Reset’ to reactivate it.

Do you have any questions?

2 MOSTLY OPERATOR CONTROL

2 Mostly Operator: After the stimulus appears, you must first identify it as a 'target' or 'non' within 3 seconds, and then the machine will provide its suggestion.

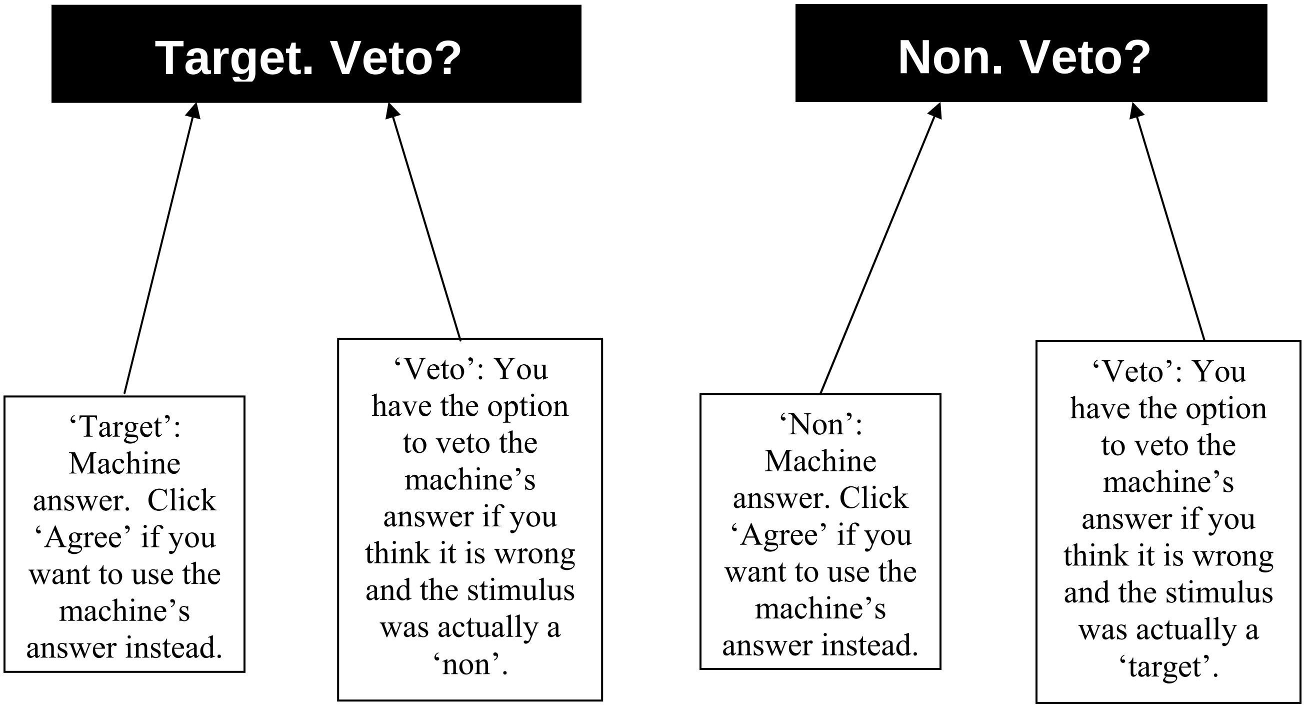
If you like your original answer and think you are correct, then ignore the machine suggestion and wait for the next stimulus. Your answer is the default answer.

However, you can **change your answer** based on the machine's suggestion, **if you wish**.

If you think the machine is correct and want the machine's suggestion to be used as your answer, click on the **'Agree' button**. If you think the machine is incorrect, you can click on the **'Veto' button** to correct the machine's identification, **or simply do nothing** and then your answer will count.

You must click on 'Agree' or 'Veto' within 3 seconds if you choose to do so.

<u>When the stimulus is a line:</u>



'Target': Machine answer. Click 'Agree' if you want to use the machine's answer instead.

'Veto': You have the option to veto the machine's answer if you think it is wrong and the stimulus was actually a 'non'.

'Non': Machine answer. Click 'Agree' if you want to use the machine's answer instead.

'Veto': You have the option to veto the machine's answer if you think it is wrong and the stimulus was actually a 'target'.

When the stimulus is a tone:

If a target tone is presented, the machine voice in the headphones will say, **"Target. Veto?"** Click 'Agree' if you want to use the machine's answer instead. Or, you have the option to veto the machine's answer if you think it is wrong and the stimulus was actually a 'non'.

If a non tone is presented, the machine voice in the headphones will say, **"Non. Veto?"** Click 'Agree' if you want to use the machine's answer instead. Or, you have the option to veto the machine's answer if you think it is wrong and the stimulus was actually a 'target'.

Clicking on 'Agree' or 'Veto' after the machine provides its answer is **always optional**. You can either respond to the machine's answers or ignore them, depending on how much you trust the machine's decisions, how much you want to interact with the system, and how much of the task you want to control. You can choose to either respond to the machine's answers or ignore them, but your task is to ensure the highest possible number of correct identifications

Do you have any questions?

You will now view a demonstration of the '2 Mostly Operator' level in the Visual Control Mode and Auditory Control Mode.

Now you will practice with the identifying the lines and tones as targets or nons at the ‘2 Mostly Operator’ level.

Visual Task Practice settings:

- **Visual Control Mode: 2 Mostly Operator**
- **Auditory Control Mode: 5 Full Machine**

Auditory Task Practice settings:

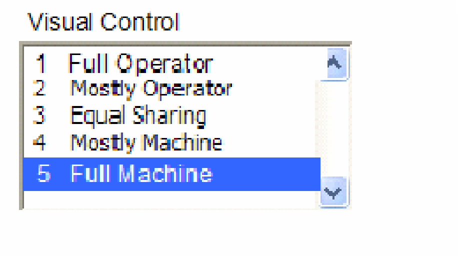
- **Visual Control Mode: 5 Full Machine**
- **Auditory Control Mode: 2 Mostly Operator**

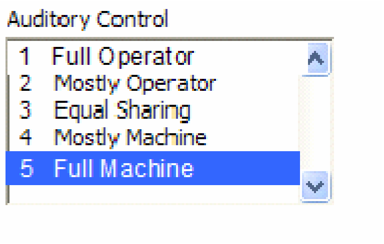
Control Modes for Machine Assistance

Visual Control Mode: The level of machine assistance for identifying the **lines**.

Auditory Control Mode: The level of machine assistance for identifying the **tones**.

The level of machine assistance which is currently engaged is highlighted on the Auditory Control and the Visual Control menus on the screen. In the example below, both the Visual Control and Auditory Control modes are set at the '5 Full Machine' level.





TEST SESSION: IDENTIFYING ONLY LINES

Now you will begin a test session, in which you are responsible for identifying the **lines only** as targets or nons.

<u>Test settings:</u>

- **Visual Control Mode: 2 Mostly Operator (you identify all lines)**
- Auditory Control Mode: Fixed at 5 Full Machine (machine will identify tones all by itself, so you can ignore the tones)

If you think the machine decision aid has stopped identifying the tones, click on ‘Reset’ to reactivate it.

Remember that the machine is **90% reliable** in its decisions. It will make some mistakes

The participant with the best overall performance will receive a **$100 cash reward**.

Do you have any questions?

TEST SESSION: IDENTIFYING ONLY TONES

Now you will begin a test session, in which you are responsible for identifying the **tones only** as targets or nons.

Test settings:

- Visual Control Mode: 5: Fixed Full Machine (the machine will identify all lines, so you can ignore the lines)
- **Auditory Control Mode: 2 Mostly Operator (you identify all tones)**

If you think the machine decision aid has stopped identifying the lines, click on ‘Reset’ to reactivate it.

Remember that the machine is **90% reliable** in its decisions. It will make some mistakes.

The participant with the best overall performance will receive a **$100 cash reward**.

Do you have any questions?

TEST SESSION: IDENTIFYING BOTH LINES AND TONES

Now you will begin the final session, in which you are responsible for identifying the both the **lines and tones** as targets or nons.

Test settings:

- **Visual Control Mode: 2 Mostly Operator (you identify all lines)**
- **Auditory Control Mode: 2 Mostly Operator (you identify all tones)**

Remember that the machine is **90% reliable** in its decisions. It will make some mistakes.

The participant with the best overall performance will receive a **$100 cash reward**.

Do you have any questions?

APPENDIX N: EXPERIMENT 3 TASK INSTRUCTIONS FOR PARTICIPANTS FIXED IN LOA 4

You are responsible for supervising a machine that identifies visual and auditory targets.

You will **see vertical lines** on the screen and **hear tones** in your headphones. Lines will appear on the left and right sides of the screen, and the tones will be presented in the left and right earphones separately. Therefore, there will be four types of stimuli, but they will never occur at the same time. They will occur in random orders.

Your task is to make sure that the machine correctly identifies each stimulus as a "Target' or 'Non' (non-target) as follows:

Visual task stimuli:

- **Target line**: presented on the screen for the **shorter** period of time (125 ms)
- **Non line**: presented on the screen for the **longer** period of time (250 ms)

Auditory task stimuli:

- **Target tone**: presented in the earphone for the **shorter** period of time (200 ms)
- **Non tone**: presented in the earphone for the **longer** period of time (250 ms)

The term 'non' is used instead of 'nontarget'.

You will first view a demonstration of how the target and non lines look.

Then you will view a demonstration of how the target and non tones sound.

The machine will identify all the stimuli immediately after they are presented.

MACHINE DECISION AID

The machine's decision aid will identify all targets and nons for you automatically. There are two different levels of machine assistance which you will use during the tasks.

They are:

5 Full Machine: Machine will make all target and non identifications automatically without any input from you

and

4 Mostly Machine: Machine will make all target and non identifications automatically, however, you can supervise its decisions and veto its answers if you wish

The machine is always **90% reliable** in its ability to make correct identifications. This means it identifies a stimulus correctly 90% of the time and incorrectly 10% of the time. It will make some mistakes.

There are three other levels of assistance which the machine can offer, but you will not be using them.

Now you will learn how to use the two levels which will be available to you, which are "5 Full Machine" and "4 Mostly Machine".

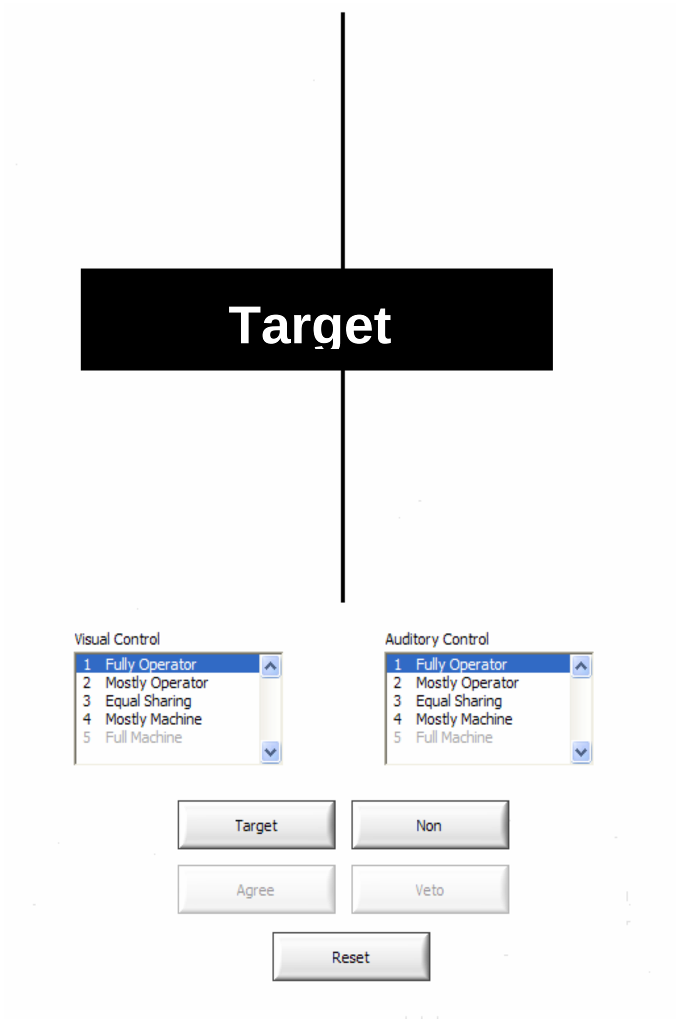
5 FULL MACHINE CONTROL

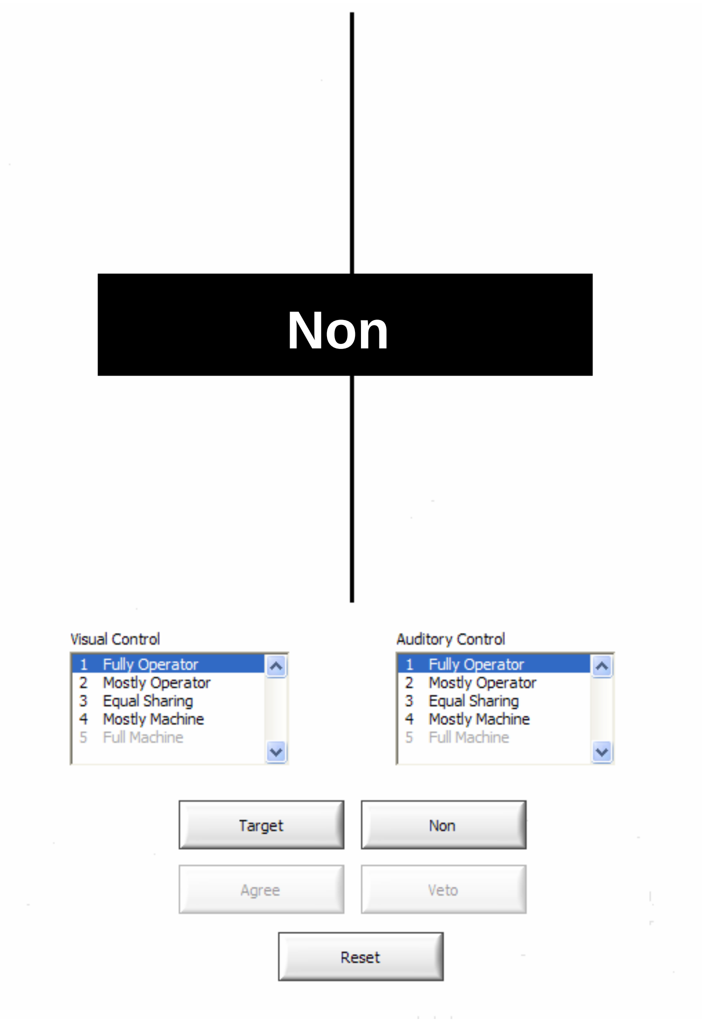
The highest level of machine assistance is ‘5 Full Machine’.

5 Full Machine: The machine makes all identifications automatically, so you do not have to do anything except make sure it still functions. You will hear its voiced answers when it identifies a tone, and you will see its message answers when it identifies a line.

The machine will provide a **message on the screen** that reads ‘Target’ after a target line appears and ‘Non’ after a non line appears in the visual task.

The machine will also provide a **voice in the headphones** that says ‘Target’ after a target tone and ‘Non’ after a non tone in the auditory task.





You are still responsible for making sure the machine is making its decisions by checking for its answer messages on the screen and voice messages in the earphones. You do not have to pay attention to the lines or tones, but if you think the machine decision aid has stopped working properly (it makes too many mistakes or quits providing answers), click on ‘Reset’ to reactivate it.

Do you have any questions?

4 MOSTLY MACHINE CONTROL

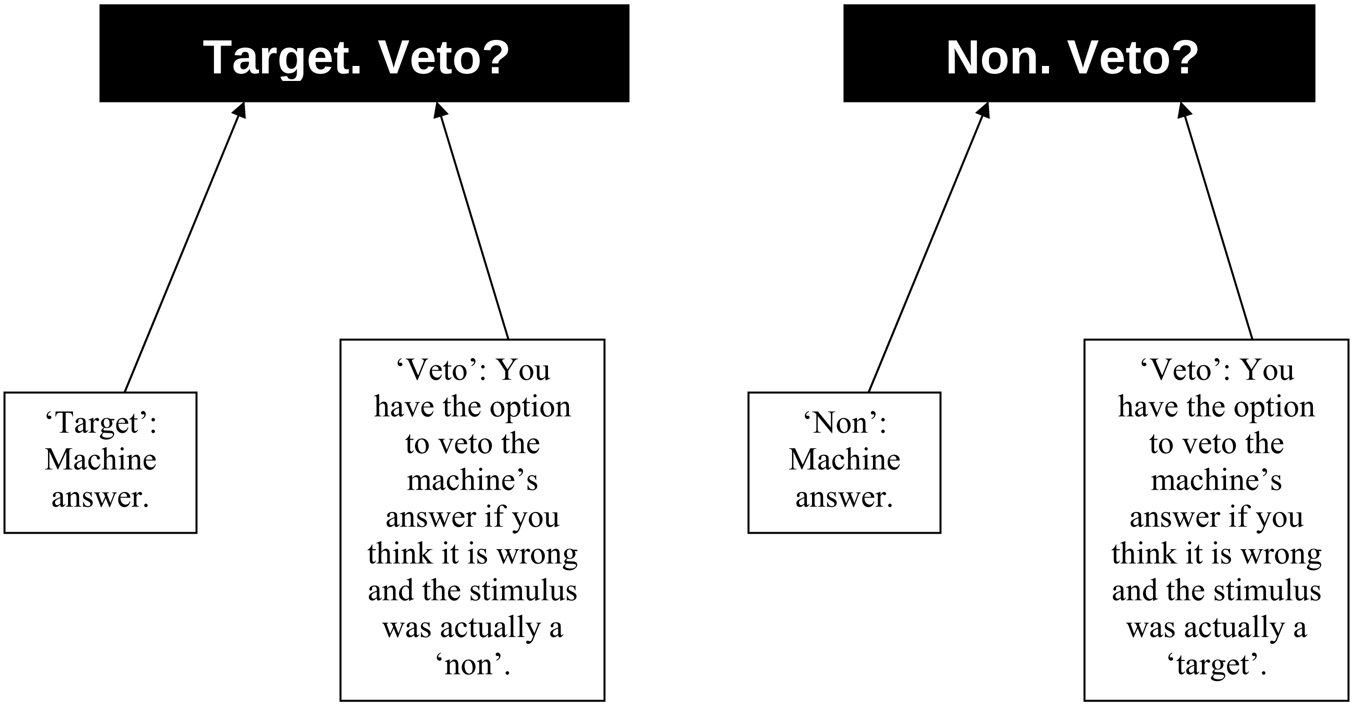
4 Mostly Machine: The machine makes all identifications automatically, but you can veto its answers if you think it has made a mistake.

The machine will identify the stimulus as a ‘target’ or ‘non’ immediately after it is presented.

If you wish, you can check the machine’s answers to make sure they are correct, because the machine can make mistakes.

If you think the machine is correct, you can either **click on ‘Agree’ or simply do nothing** and then the machine’s answer will count as your answer. If you think the machine is incorrect, you can click on the **‘Veto’** button to correct its suggestion. If you do not click on “Agree’ or “Veto’, the **machine assumes you agree** with it, and its answer will count towards your score.
You must click on ‘Agree’ or ‘Veto’ within 3 seconds if you choose to do so.

When the stimulus is a line:



‘Target’: Machine answer.

‘Veto’: You have the option to veto the machine’s answer if you think it is wrong and the stimulus was actually a ‘non’.

‘Non’: Machine answer.

‘Veto’: You have the option to veto the machine’s answer if you think it is wrong and the stimulus was actually a ‘target’.

When the stimulus is a tone:

If a target tone is presented, the machine voice in the headphones will say, **“Target. Veto?”** You have the option to veto the machine’s answer if you think it is wrong and the stimulus was actually a ‘non’.

If a non tone is presented, the machine voice in the headphones will say, **“Non. Veto?”** You have the option to veto the machine’s answer if you think it is wrong and the stimulus was actually a ‘target’.

Clicking on ‘Agree’ or ‘Veto’ after the machine provides its answer is **always optional**. You can either respond to the machine’s answers or ignore them, depending on how much you trust the machine’s decisions, how much you want to interact with the system, and how much of the task you want to control. You can choose to either respond to the machine’s answers or ignore them, but your task is to ensure the highest possible number of correct identifications.

Do you have any questions?

You will now view a demonstration of the ‘4 Mostly Machine’ level in the Visual Control Mode and Auditory Control Mode.

Now you will practice with the identifying the lines and tones as targets or nons at the '4 Mostly Machine' level.

<u>Visual Task Practice settings:</u>

- **Visual Control Mode: 4 Mostly Machine**
- **Auditory Control Mode: 5 Full Machine**

<u>Auditory Task Practice settings:</u>

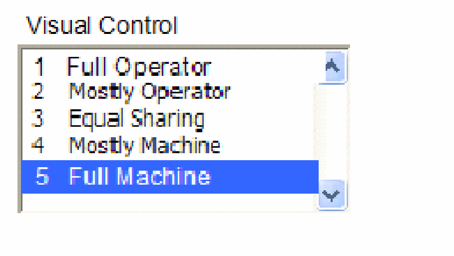
- **Visual Control Mode: 5 Full Machine**
- **Auditory Control Mode: 4 Full Machine**

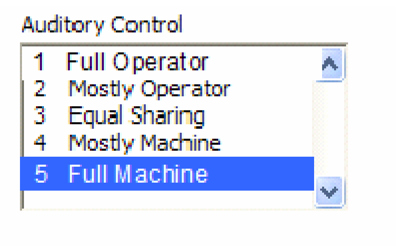
Control Modes for Machine Assistance

Visual Control Mode: The level of machine assistance for identifying the **lines**.

Auditory Control Mode: The level of machine assistance for identifying the **tones**.

The level of machine assistance which is currently engaged is highlighted on the Auditory Control and the Visual Control menus on the screen. In the example below, both the Visual Control and Auditory Control modes are set at the '5 Full Machine' level.





TEST SESSION: IDENTIFYING ONLY LINES

Now you will begin a test session, in which you are responsible for making sure the machine correctly identifies the **lines only** as targets or nons.

Settings:

- **Visual Control Mode: 4 Mostly Machine (you supervise identification of all lines)**
- Auditory Control Mode: Fixed at 5 Full Machine (machine will identify tones all by itself, so you can ignore the tones)

If you think the machine decision aid has stopped identifying the tones, click on ‘Reset’ to reactivate it.

Remember that the machine is **90% reliable** in its decisions. It will make some mistakes.

The participant with the best overall performance will receive a **$100 cash reward**.

TEST SESSION: IDENTIFYING ONLY TONES

Now you will begin a test session, in which you are responsible for making sure the machine correctly identifies the **tones only** as targets or nons.

Settings:

- Visual Control Mode: 5: Fixed Full Machine (the machine will identify all lines, so you can ignore the lines)
- **Auditory Control Mode: 4 Mostly Machine (you supervise identification of all tones)**

If you think the machine decision aid has stopped identifying the lines, click on 'Reset' to reactivate it.

Remember that the machine is **90% reliable** in its decisions. It will make some mistakes.

The participant with the best overall performance will receive a **$100 cash reward**.

TEST SESSION: IDENTIFYING BOTH LINES AND TONES

Now you will begin the final session, in which are responsible for making sure the machine correctly identifies both lines and tones.

Test settings:

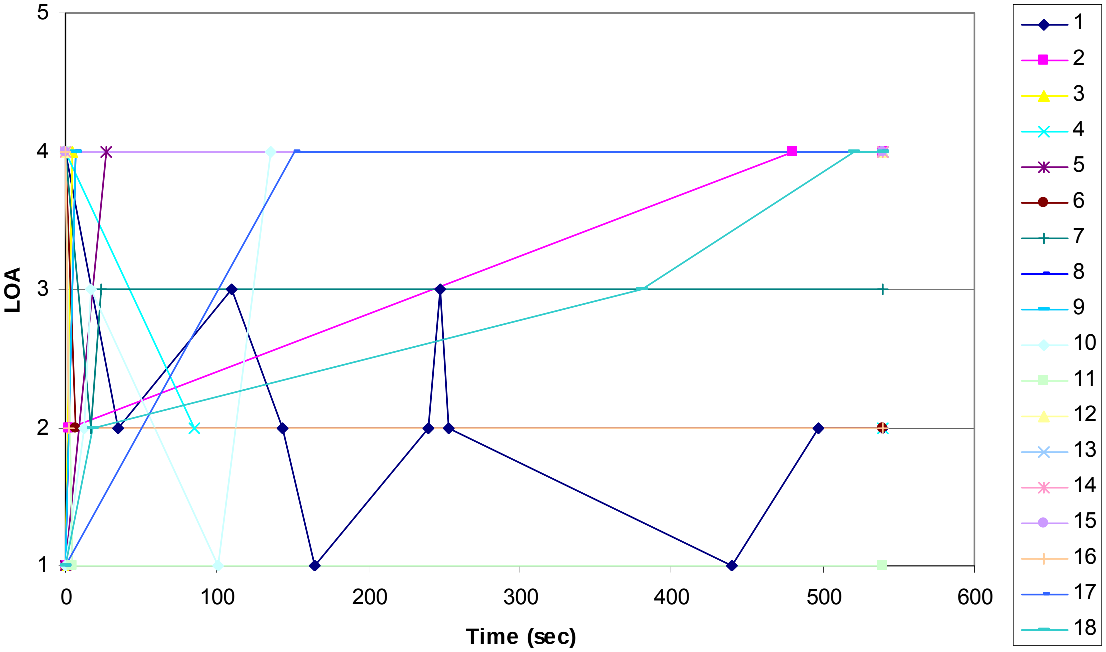
- **Visual Control Mode: 4 Mostly Machine (you supervise identification of all lines)**

- **Auditory Control Mode: 4 Mostly Machine (you supervise identification of all tones)**

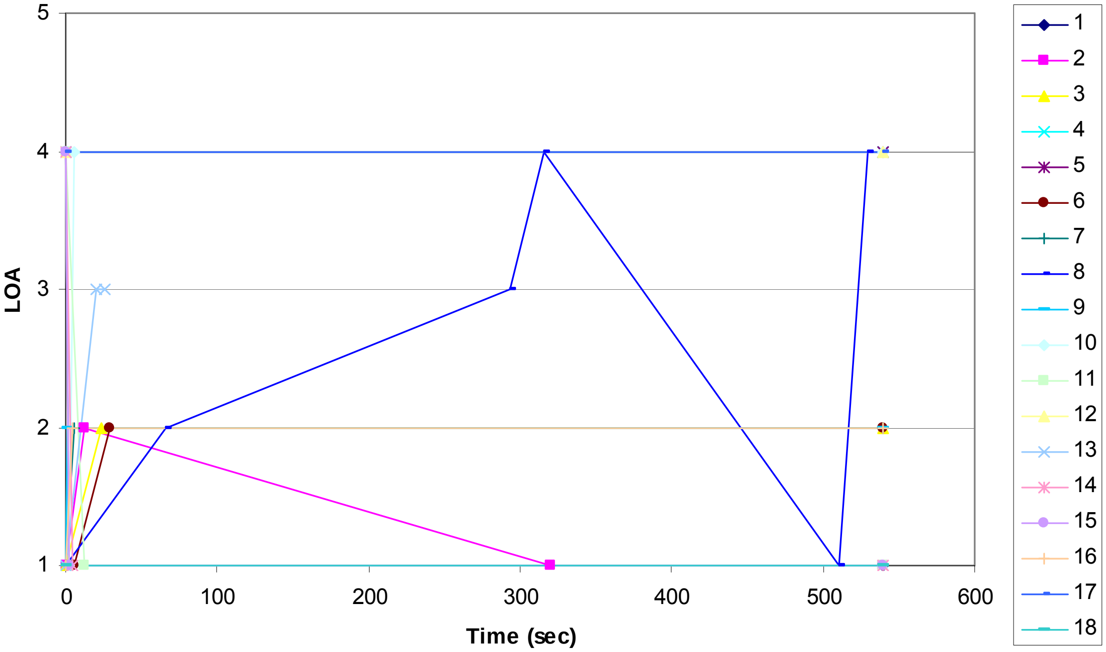
Remember that the machine is **90% reliable** in its decisions. It will make some mistakes.

The participant with the best overall performance will receive a **$100 cash reward**.

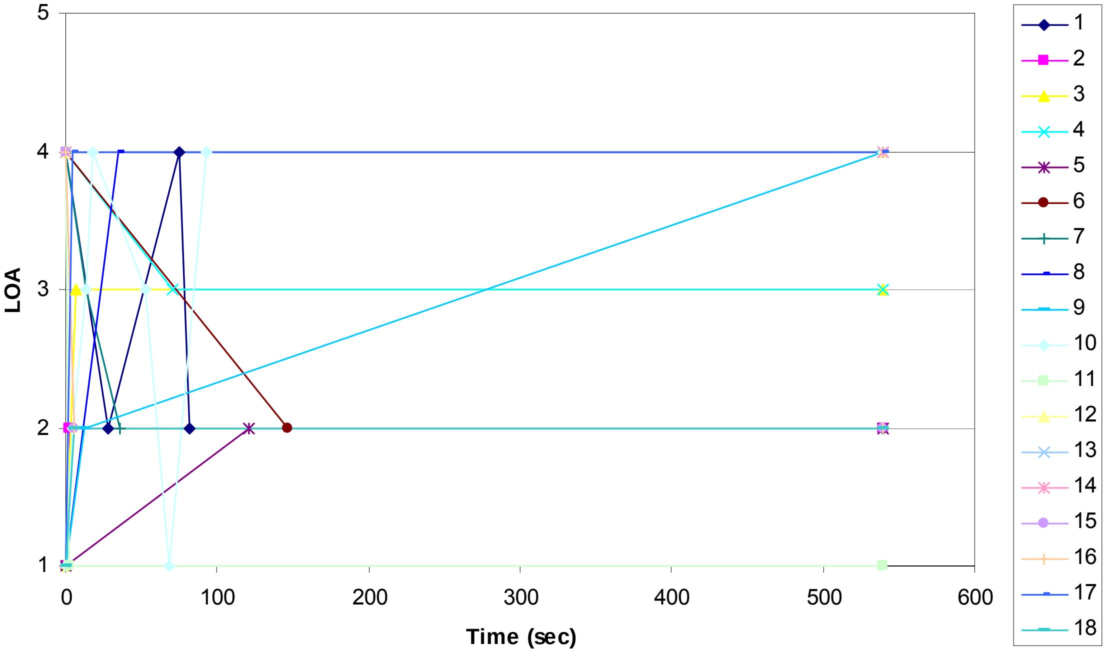
APPENDIX O: DUTY CYCLE OF LOW AC PARTICIPANTS IN THE VISUAL-ONLY TASK


LOA
5
4
3
2
1
0
100
200
300
400
500
600
Time (sec)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

APPENDIX P: DUTY CYCLE OF HIGH AC PARTICIPANTS IN THE VISUAL-ONLY TASK

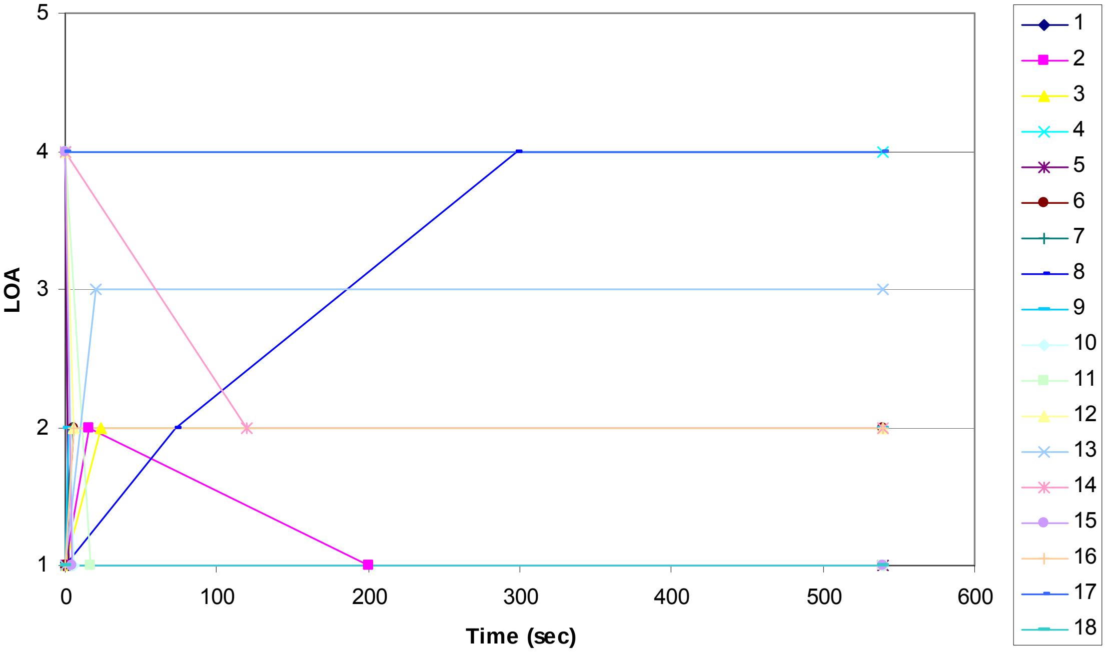

LOA
5
4
3
2
1
Time (sec)
0
100
200
300
400
500
600
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

APPENDIX Q: DUTY CYCLE OF LOW AC PARTICIPANTS IN THE AUDITORY-ONLY TASK

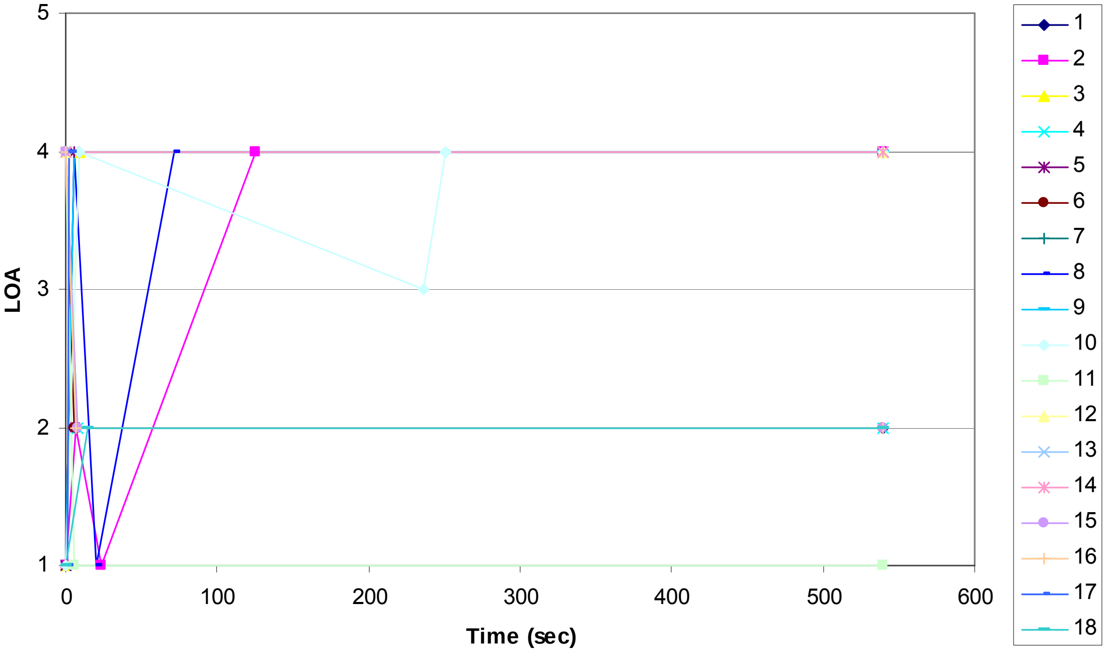


LOA
5
4
3
2
1
0
100
200
300
400
500
600
Time (sec)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

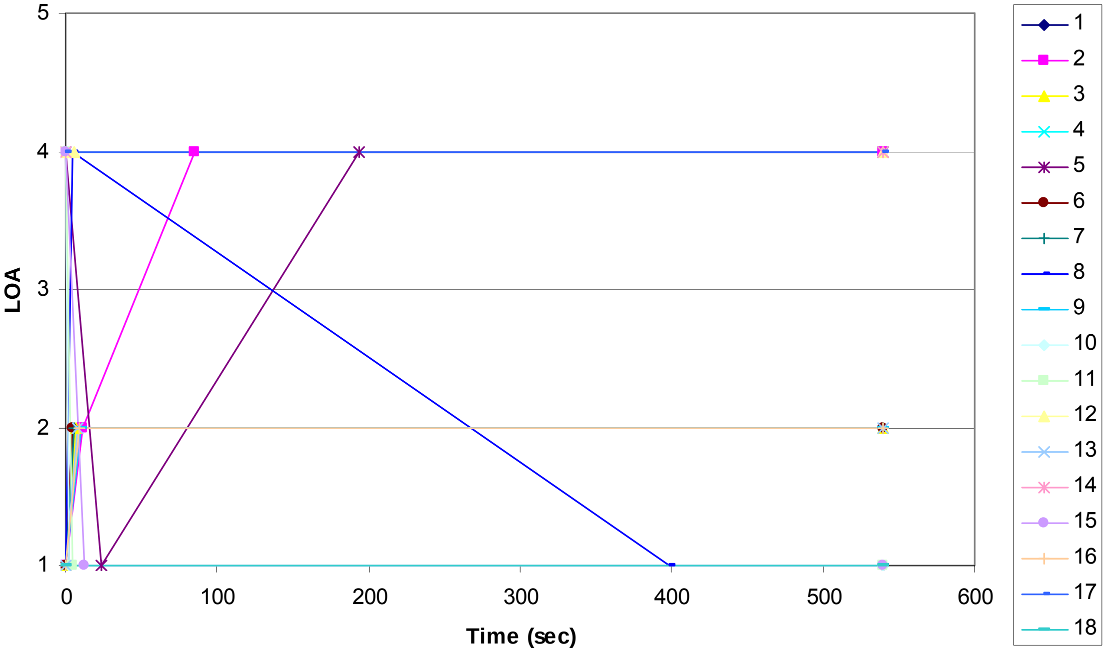
APPENDIX R: DUTY CYCLE OF HIGH AC PARTICIPANTS IN THE AUDITORY-ONLY TASK


LOA
5
4
3
2
1
0
100
200
300
400
500
600
Time (sec)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

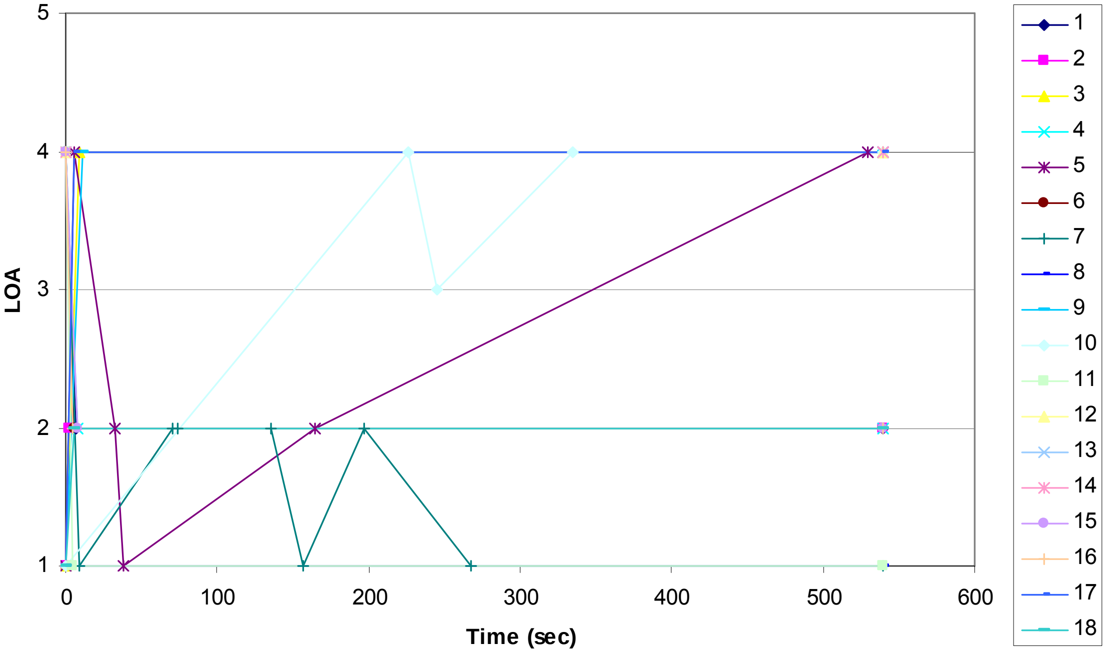
APPENDIX S: DUTY CYCLE OF LOW AC PARTICIPANTS IN THE VISUAL-COMBINED TASK


5
4
3
2
1
LOA
0
100
200
300
400
500
600
Time (sec)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

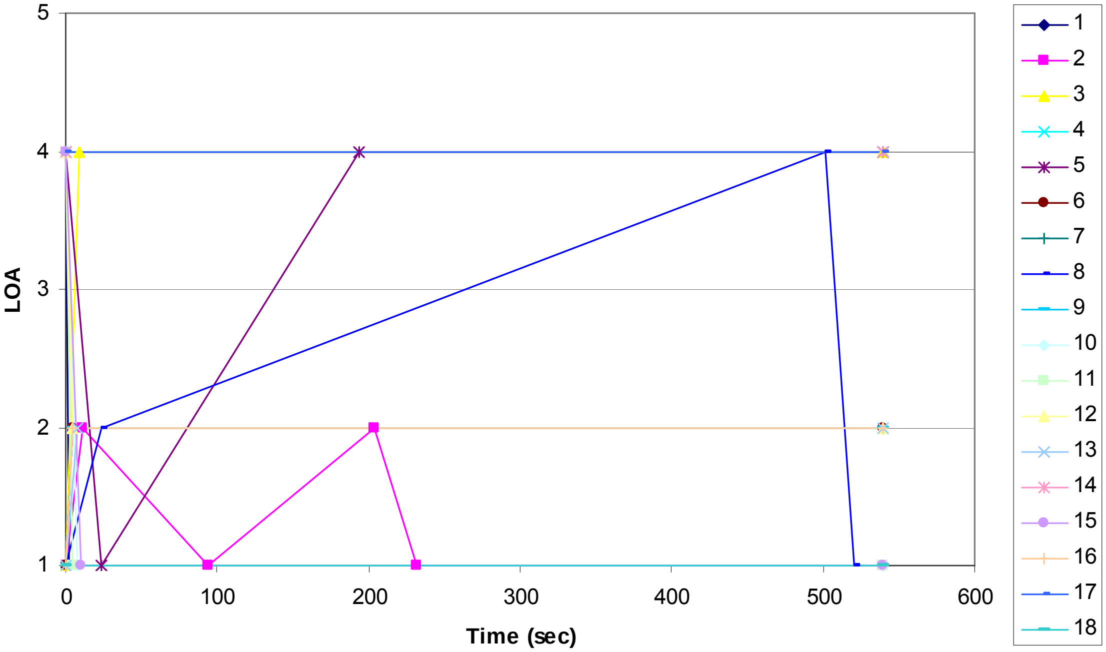
APPENDIX T: DUTY CYCLE OF HIGH AC PARTICIPANTS IN THE VISUAL-COMBINED TASK


5
4
3
2
1
LOA
0
100
200
300
400
500
600
Time (sec)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

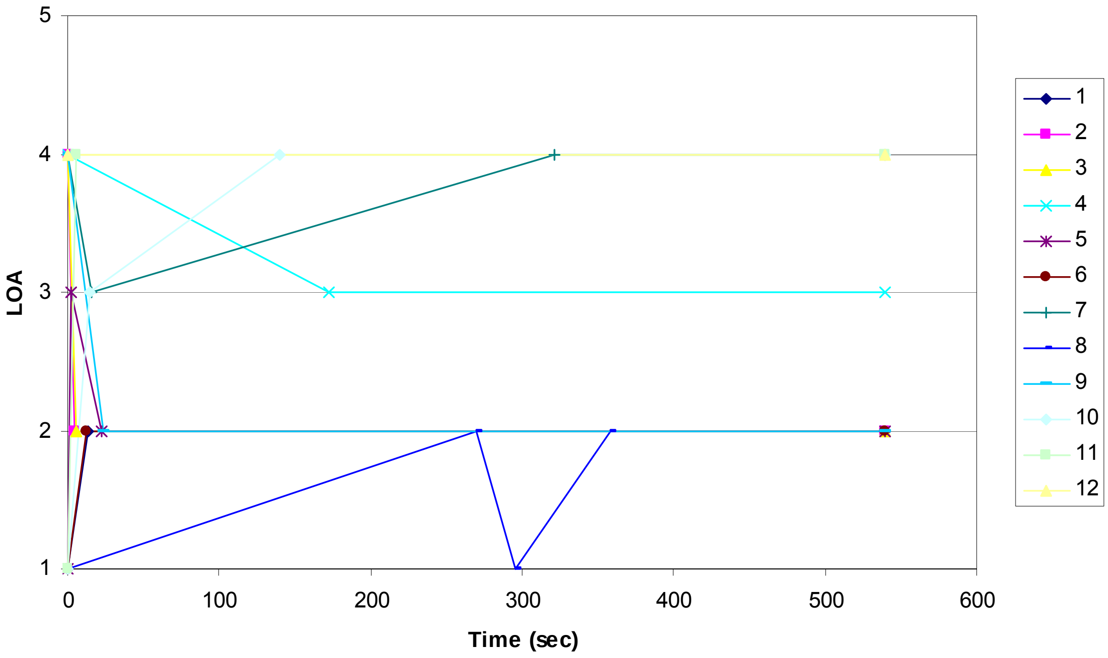
APPENDIX U: DUTY CYCLE OF LOW AC PARTICIPANTS IN THE AUDITORY-COMBINED TASK


5
4
3
2
1
LOA
0
100
200
300
400
500
600
Time (sec)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

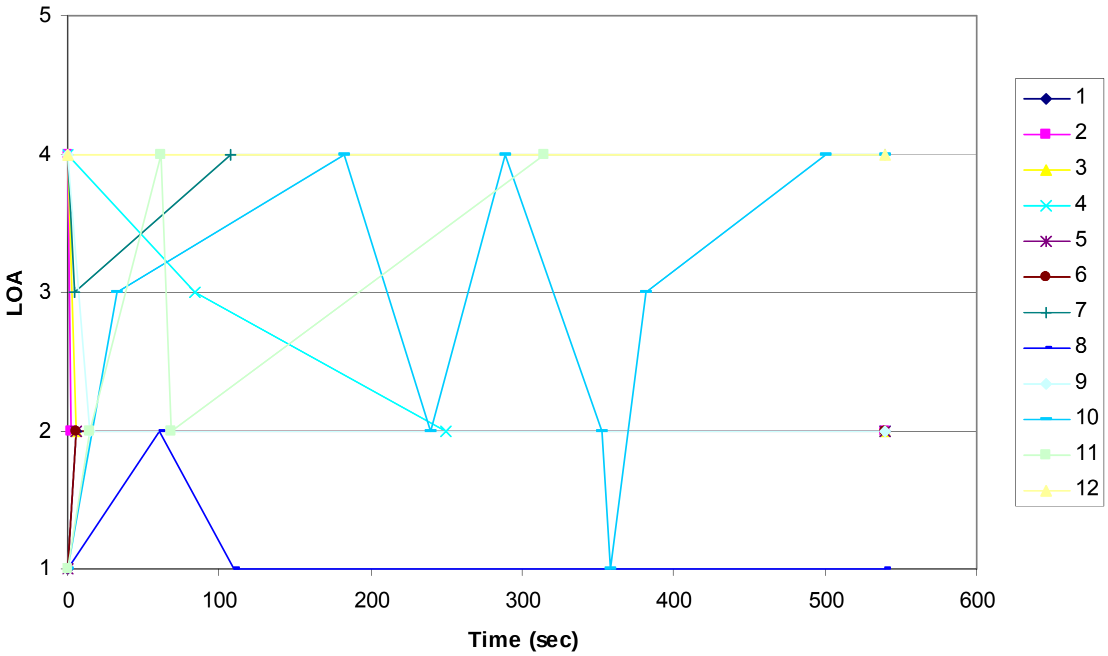
APPENDIX V: DUTY CYCLE OF HIGH AC PARTICIPANTS IN THE AUDITORY-COMBINED TASK


5
4
3
2
1
LOA
0
100
200
300
400
500
600
Time (sec)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

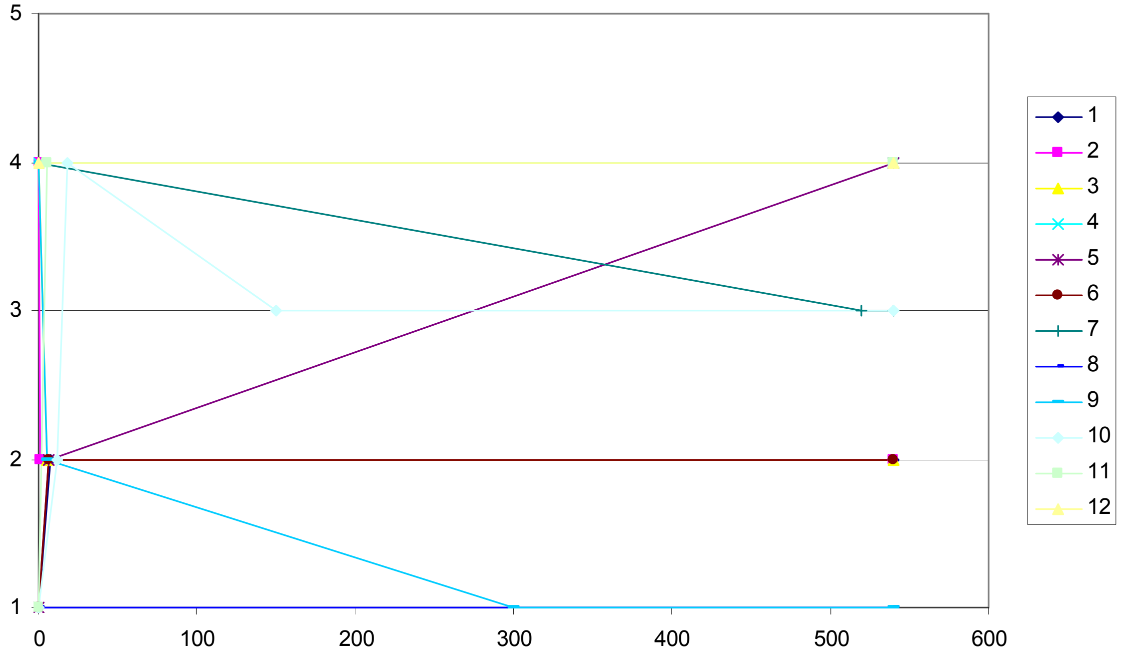
APPENDIX W: DUTY CYCLE OF MEDIUM AC PARTICIPANTS IN THE VISUAL-ONLY TASK


5
4
3
2
1
LOA
0
100
200
300
400
500
600
Time (sec)
1
2
3
4
5
6
7
8
9
10
11
12

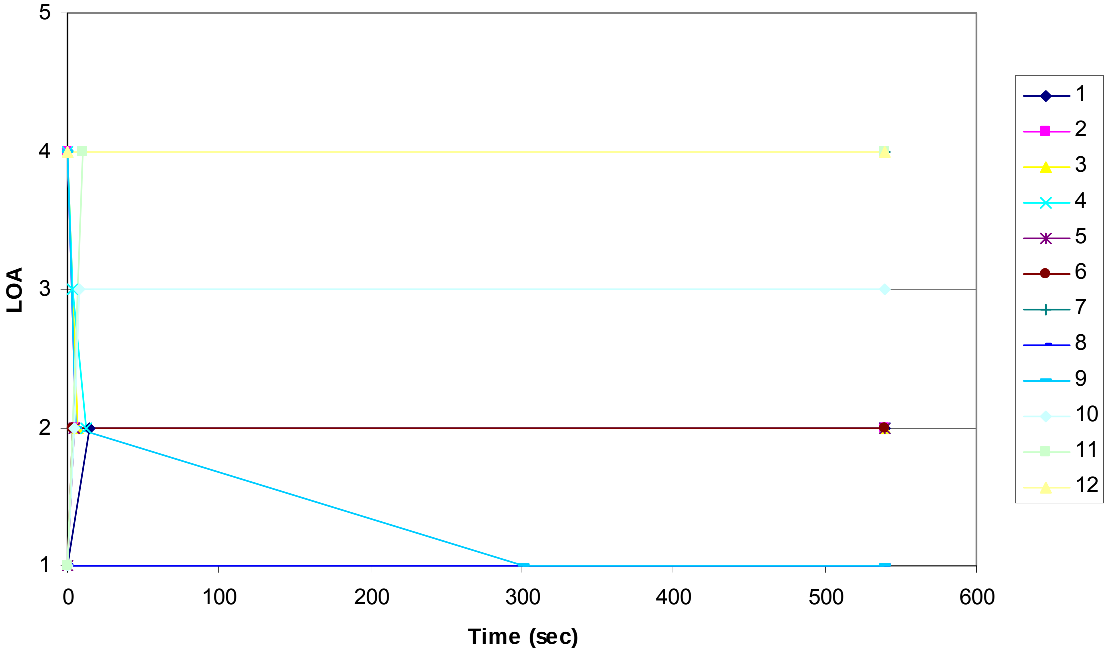
APPENDIX X: DUTY CYCLE OF MEDIUM AC PARTICIPANTS IN THE AUDITORY-ONLY TASK


5
4
3
2
1
LOA
0
100
200
300
400
500
600
Time (sec)
1
2
3
4
5
6
7
8
9
10
11
12

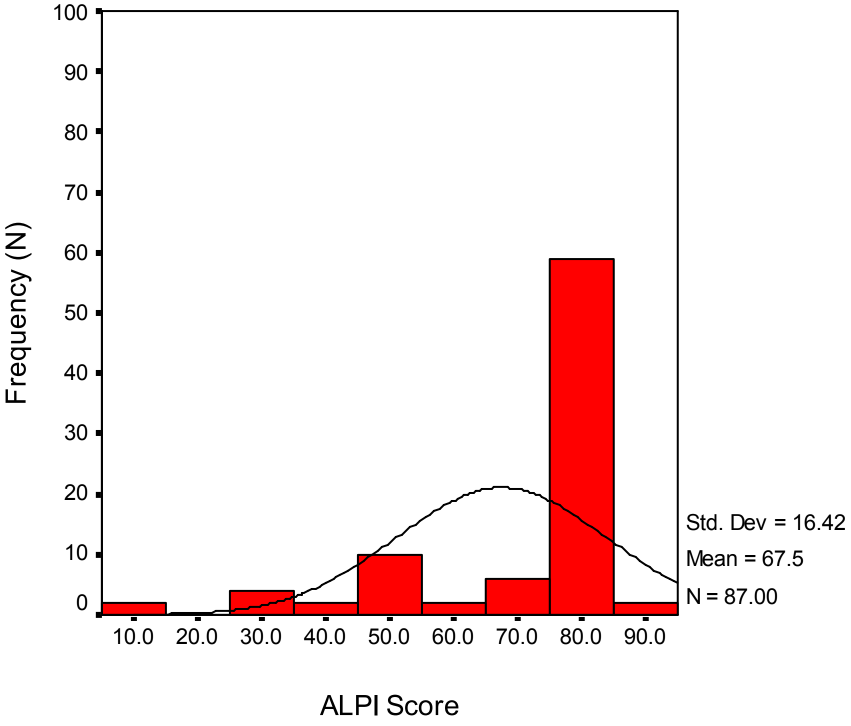
APPENDIX Y: DUTY CYCLE OF MEDIUM AC PARTICIPANTS IN THE VISUAL-COMBINED TASK


5
4
3
2
1
0
100
200
300
400
500
600
1
2
3
4
5
6
7
8
9
10
11
12

APPENDIX Z: DUTY CYCLE OF MEDIUM AC PARTICIPANTS IN THE AUDITORY-COMBINED TASK

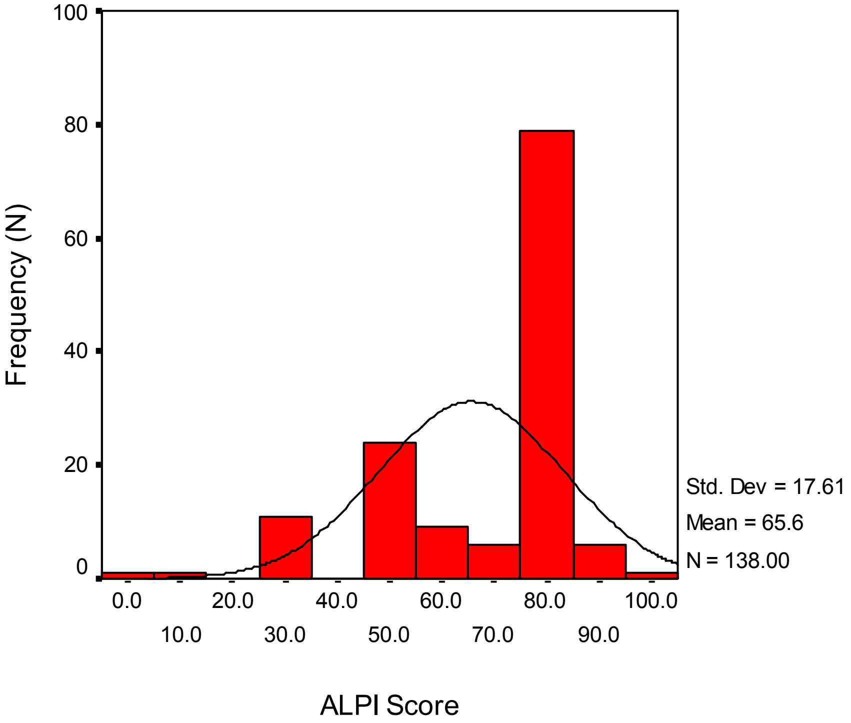

5
4
3
2
1
LOA
0
100
200
300
400
500
600
Time (sec)
1
2
3
4
5
6
7
8
9
10
11
12

APPENDIX AA: DISTRIBUTION OF ALPI SCORES IN THE LOW AC GROUP IN EXPERIMENT 1

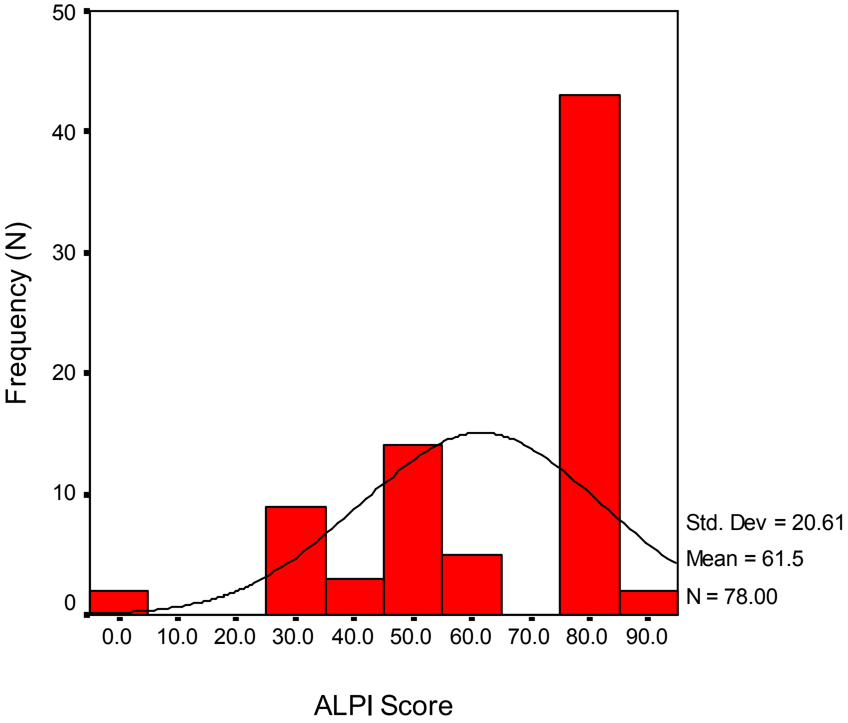


100
90
80
70
60
50
40
30
20
10
0
Frequency (N)
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0
90.0
ALPI Score
Std. Dev = 16.42
Mean = 67.5
N = 87.00

APPENDIX AB: DISTRIBUTION OF ALPI SCORES IN THE MEDIUM AC GROUP IN EXPERIMENT 1


100
80
60
40
20
0
Frequency (N)
Std. Dev = 17.61
Mean = 65.6
N = 138.00
0.0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0
90.0
100.0
ALPI Score

APPENDIX AC: DISTRIBUTION OF ALPI SCORES IN THE HIGH AC GROUP IN EXPERIMENT 1


50
40
30
20
10
0
Frequency (N)
0.0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0
90.0
ALPI Score
Std. Dev = 20.61
Mean = 61.5
N = 78.00

REFERENCES

Awh, E., Jonides, J., & Reuter-Lorenz, P.A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception & Performance, 24,* 780-790.

Baddeley, A.D. (1995). Working memory. In M.S. Gazzaniga et al. (Eds.), *The cognitive neurosciences*. (pp. 775-784). Cambridge, MA: MIT Press.

Bartram, D. & Dale, H. (1982). The Eysenck Personality Inventory as a selection test for military pilots. *Journal of Occupational Psychology, 55*, 287-296.

Bellenkes, A.H., Wickens, C.D., & Kramer, A.F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine, 68,* 569-579.

Berkun, M.M. (1964). Performance decrement under psychological stress. *Human Factors, 6,* 21-30.

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

Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21,* 1-16.

Bubb-Lewis, C. & Scerbo, M.W. (1999). Does desire for control affect interactions in an adaptive automated environment? In M. W. Scerbo & M. Mouloua (Eds.), *Automation Technology and Human Performance* (pp. 124-128). Mahwah, NJ: Lawrence Erlbaum Associates.

Burger, J.M. & Cooper, H.M. (1979). The desirability of control. *Motivation and Emotion*, *3*, 381-393.

Calvo, M. G. & Eysenck, M. W. (1998). Cognitive bias to internal sourced of information in anxiety. *International Journal of Psychology,* 33, 287-299.

Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America,25,* 975-979.

Costa, P.T. & McCrae, R.R. (1992a). *NEO PI-R Professional Manual.* Odessa, FL: Psychological Assessment Manuals.

Craig, A., Davies, D.R., & Matthews, G. (1987). Diurnal variation, task characteristics and vigilance performance. *Human Factors, 29,* 675-684.

Daneman, M., & Tardif, T. (1987). Working memory and reading skill re-examined. In M. Coltheart, Attention and performance, Vol. XII (pp. 491-508). Hillsdale: LEA.

Darke, S. (1988). Anxiety and working memory capacity. *Cognition and Emotion*, *2*, 145-154.

Darwin, C., Turvey, M.T., & Crowder, R.G. (1972). An analog of the Sperling partial report procedure. *Cognitive Psychology, 3,* 255-267.

Davies, D.R. & Parasuraman, R. (1982). *The Psychology of Vigilance.* London: Academic Press.

Dember, W.N. & Warm, J.S. (1979). *Psychology of perception* (2nd ed.). New York: Holt, Rinehart, and Winston.

Derryberry, D. & Reed, M.A. “Anxiety-related attentional biases and their regulation by attentional control,” *Journal of Abnormal Psychology*, vol. 111, no. 2, pp. 225-236, 2002.

Dorneich, M.C., Whitlow, S.D., Ververs, P.M., & Rogers, W.H. (2003). Mitigating cognitive bottlenecks via an augmented cognition adaptive system. *IEEE International Conference on Systems, Man, and Cybernetics,* , pp. 937-944.

Droit-Volet, S., Tourret, S. & Wearden, J. (2004). Perception of the duration of auditory and visual stimuli in children and adults. *The Quarterly Journal of Experimental Psychology,57A*(5), 797-818.

Edmonds, E.A. (1981). Adaptive man-computer interfaces. In M. J. Coombs & J.L. Alty (Eds.), *Computing Skills and the User Interface,*(pp. 389-426). London: Academic Press, 1981.

Edwards, E. (1988). Introductory overview. In E. Wiener & D. Nagel (Eds.), *Human Factors in Aviation,* (pp. 3-25). San Diego, CA: Academic Press.

Egan, D.E. (1988). Individual differences in human-computer interaction. In M. Helander (Ed.), *Handbook of Human-Computer Interaction,* Amsterdam: North Holland.

Egan, D. E. (1988). Individual differences in human-computer interaction. In Handbook of Human-Computer Interaction (M. Helander, Ed.) Elsevier Science Publishers B.V. (North-Holland).

Egan, J., Carterette, E., & Thwing, E. (1954). Some factors affecting multichannel listening. *Journal of the Acoustical Society of America, 26,* 774-782.

Elliman, N. A., Green, M. W., Rogers, P. J., & Finch, G. M. (1996). Processing-efficiency theory and the working memory system: Impairments associated with sub-clinical anxiety. *Personality and Individual Differences, 23*(1), 31-35.

Endsley, M.R. & Bolstad, C.A. (1994). Individual differences in pilot situation awareness. *International Journal of Aviation Psychology*, *4*(3), 241-264.

Endsley, M.R. (1995). A taxonomy of situation awareness errors. In R. Fuller, N. Johnston, & N. McDonald (Eds.), *Human Factors in Aviation Operations* (pp. 287-292)*,* Aldershot, England: Avebury Aviation, Ashgate Publishing Ltd.

Engle, R. W. (2002). Working memory capacity as executive function. *Current Directions in Psychological Science, 11*(1), 19-23.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999a). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake, & P. Shah (Eds.), *Models of working memory* (pp. 102-134). Cambridge: Cambridge University Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211-245.

Eysenck, H.J. & Eysenck, M.W. (1985). *Personality and individual differences: A natural science approach.* New York: Plenum.

Eysenck, M. W. & Derakshan, N. (1997). Cognitive biases for future negative events as a function of trait anxiety and social desirability. *Personality and Individual Differences, 22*, 597-605.

Eysenck, M.W. (1981). Learning, memory, and personality. In H. J. Eysenck (Ed.), *A model for personality* (pp. 169-209). New York: Springer-Verlag.

Eysenck, M.W. (1992). *Anxiety: The cognitive perspective.* Hove, UK: Lawrence Erlbaum Associated. Ltd.

Eysenck, M.W. (1997). *Anxiety and cognition: A unified theory.* Hove, UK: Psychology Press.

Field, A. (2000). *Discovering statistics using SPSS for Windows*. London: Sage Publications, Ltd.

Fitts, P.M. (1951). *Human Engineering for an Effective Air Navigation and Traffic Control System* (Selected extract). Washington, DC: National Research Council.

Funk, K.H. II & McCoy, B. (1996). A functional model of flight deck agenda management. In *Proceedings of the 40th annual meeting of the Human Factors and Ergonomics Society* (pp. 254-258). Santa Monica, CA: Human Factors and Ergonomics Society.

Gescheider, G.A. (1997). *Psychophysics: The fundamentals* (3rd ed.). Mahwah, NJ: Erlbaum.

Gopher, D. (1991). The skill of attentional control: Acquisition and execution of attention atrategies. In D. Meyer & S. Kornblum (Eds.), *Attention and Performance IVX.* Hillsdale, NJ: Erlbaum.

Gopher, D., Weil, M., & Seigel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica, 71,* 147-177.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Gugerty, L. & Tirre, W. (1997). Situation awareness: A validation study and investigation of individual differences. In *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting* (pp. 564-568). Santa Monica, CA: Human Factors and Ergonomics Society.

Hammer, J. (1999). Human factors of functionality in avionics. In D. Garland, J. Wise, & V.D. Hopkin (Eds.), *Handbook of aviation human factors.* Mahwah, NJ: Erlbaum.

Hammer, J.M. &. Small, R.L. (1995). An intelligent interface in an associate system. In W. B. Rouse (Ed.), *Human-Technology Interaction in Complex Systems,* vol. 7, (pp. 1-44). Greenwich, CT: JAI Press.

Hancock, P.A. & Chignell, M.H. (1989). Adaptive control in human-machine systems. In P.A. Hancock (Ed.), *Human Factors Psychology,* (pp. 305-345). Amsterdam: North-Holland Publishing Co.

Hancock, P.A. & Scallen, S.F. (1996). The future of function allocation. *Ergonomics in Design,* 4(4), 24-29.

Hancock, P.A. The effects of automation invocation procedure and dynamic display relocation on performance in a multi-task environment. Manuscript submitted.

Hancock. P. A. & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors, 31(5),* 519-537.

Harris, R.L. & Christhilf, D.M. (1980). What do pilots see in displays? In G. Corrick, E. Hazeltine, & R. Durst (Eds.), *Proceedings of the 24th Annual Meeting of the Human Factors and Ergonomics Society.* Santa Monica, CA: Human Factors Society.

Harris, W.C., Hancock, P.A., & Arthur, E.J. (1993). The effect of taskload projection on automation use, performance, and workload. In *Proceedings of the 7th Annual International Symposium on Aviation Psychology*.

Harris, W.C., Hancock, P.A., Arthur, E., & Caird, J.K. (1991). Automation Influences on Performance, Workload, and Fatigue (Tech. Rep. 91-N01), Human Factors Research Laboratory, University of Minnesota, Minneapolis, MN.

Hart, S. G. & Staveland, L. E. (1988). Development of a NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload.* Amsterdam: New Holland.

Hart, S.G. & Wickens, C.D. (1990). Workload assessment and prediction. In H.R. Booher (Ed.), *MANPRINT: An emerging technology. Advanced concepts for integrating people, machines, and organizations* (pp. 257-300). New York: Van Nostrand Reinhold.

Hatfield, J.L. & Loeb, M. (1968). Sense mode and coupling in a vigilance task. *Perception and Psychophysics, 4,* 29-36.

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

Hilburn, B., Molloy, R., Wong, D., & Parasuraman, R. (1993). Operator versus computer control of adaptive automation. In J. G. Morrison, Ed. Warminster (Eds.), *The Adaptive Function Allocation for Intelligent Cockpits (AFAIC) Program: Interim Research and Guidelines for the Application of Adaptive Automation* (Tech. Rep. No. NAWCADWAR-93031-60 pp. 19-24), PA: Naval Air Warfare Center.

Highhouse, S. (2002). Assessing the candidate as a whole: A historical and critical analysis of individual psychological assessment. *Personnel Psychology, 55,* 363-396.

Hockey, G.R.J. (1986). A state control theory of adaptation to stress and individual differences in stress management. In G.R.J. Hockey, A.W.K. Gaillard, & M.G.H. Coles (Eds.), *Energetics and human information processing.* Dordrecht: Martinus Nijhoff.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Kantowitz, B.H. & Casper, P.A. (1988). Human workload in aviation. In E. Wiener & D. Nagel (Eds.), *Human Factors in Aviation* (pp. 157-187). San Diego, CA: Academic Press.

Kramer, A.F. (1991). Physiological metrics of mental workload: A review of recent progress. In D.L. Damos (Ed.), *Multiple-task performance.* London: Taylor & Francis.

Kramer, A.F., Larish, J.F., & Strayer, D.L. (1995). Training for attentional control in dual task settings: A comparison of young and old adults. *Journal of Applied Psychology: Applied, 1*(1), 50-76.

Kundel H.L. & LaFollette, P.S. (1972). Visual search patterns and experience with radiological images. *Radiology, 103,* 523-528.

L.E. McCutcheon, “The desirability of control scale: Still reliable and valid twenty years later,” *Current Research in Social Psychology,* vol. 5, no. 15, pp. 225-235, 2000.

Landauer, T.K. (1995). *The Trouble with Computers.* Cambridge, MA: MIT Press.

Lee, J.D. & Moray, N. (1992). Trust, control strategies, and allocation of function in human-machine systems. *Ergonomics, 35*, 1243-1270.

Lee, J.D. & See, K.A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*(1), 50-80.

Liu, Y. & Wickens, C.D. (1992). Use of computer graphics and cluster analysis in aiding relational judgment. *Human Factors, 34,* 165-178.

Mackworth, N.H. (1976). Ways of recording line sight. In R.A. Monty & J.W. Senders (Eds.) *Eye movements and psychological processing* (pp. 173-178). Hillsdale, NJ: Erlbaum.

MacLeod, C. & Donnellan, A.M. (1993). Individual differences in anxiety and the restriction of working memory capacity. *Personality and Individual Differences, 15*, 163-173.

Macmillian, N. A., & Creelman, C. D. (2005). Detection theory: A user's guide (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Macmillan, N.A., & Creelman, C.D. (1991). *Detection Theory: A User's Guide.* Cambridge: Cambridge University Press.

Madni, A. (1988). The role of human factors in expert systems design and acceptance. *Human Factors, 30,* 395-414.

Mason, M.V. & Thomas, R.C. (1984). Experimental adaptive interface. *Information Technology: Research, Design, Applications, 3*(3), 162-167.

Massaro D.W. & Warner, D.S. (1971). Dividing attention between auditory and visual perception. *Perception and Psychophysics, 21,* 579-574.

Matthews, G. & Deary, I.J. (1998). *Personality Traits.* New York: Cambridge University Press.

Matthews, G. Joyner, L., Gilliland, K., Campbell, S. Huggins, J., & Falconer, S. (1999).

Matthews, G., Davies, D.R., Westerman, S.J., & Stammers, R.B. (2000). *Human performance: Cognition, stress, and individual differences.* Philadelphia: Taylor and Francis.

Molloy, R. & Parasuraman, R. (1997). Pilots who rely too much on automation may not detect malfunctions. *Human Factors and Ergonomics Society*.

Moray, N. (1986). Monitoring behavior and supervising control. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), *Handbook of perception and human performance*. New York: Wiley.

Morrison, N.M. & Rouse, W.B. (1986). Adaptive Aiding for Human-Computer Control: Experimental Studies of Dynamic Task Allocation. (Tech. Rep. AAMRL-TR-86-005). Naval Air Warfare Center, Warminster, PA.

Mouloua, M., Gilson, R., & Hancock, P.A. (2003). Design controls for future unmanned aerial vehicles. *Ergonomics in Design, 11*(4), 6 -11.

Mosier, Skitka, & Korte (1994).

Mueller, J. H. (1992). Anxiety and performance. In A. P. Smith & D. M. Jones (Eds.), *Handbook of human performance. Vol. 3: State and trait.* London: Academic Press.

Muir, B.M. (1987). Trust between humans and machines, and the design of decision aides. *International Journal of Man-Machine Studies, 27*, 527-539.

Mulder, G. (1986). The concept and measurement of mental effort. In G.R.J. Hockey, A.W.K. Gaillard, & M.G.H. Coles (Eds.), *Energetics and human information processing.* Dordrecht: Martinus Nijhoff.

Navon, D. & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review, 86,* 214-255.

Norcio, A.F. & Stanley, J. (1989). Adaptive human-computer interfaces: A literature survey and perspective. *IEEE Transactions on Systems, Man and Cybernetics, 19*, 399-408.

O' Hare, D. (1997). Cognitive ability determinants of elite pilot performance. *Human Factors, 39*(4), 540-552.

O'Donnell, R.D. & Eggemeier, F.T. (1986). Workload assessment methodology. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and performance* (vol. 2). New York: Plenum.

Ohman, A. (1996). Preferential preattentive processing of threat in anxiety: Preparedness and attentional biases. In R. M. Rapee (Ed.), *Current controversies in the anxiety disorders* (pp. 253-290). New York: Guilford.

Olson, W.A. & Sarter, N.B. (2000). Automation management strategies: Pilot preferences and operational experiences. *The International Journal of Aviation Psychology*, *10,*(4), 327-341.

Parasuraman, R. (1987). Human-computer monitoring. *Human Factors Special Issue: Vigilance: Basic and Applied Research, 29*(6), 695-706.

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

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

Parasuraman, R. (1986). Vigilance, monitoring, and search. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of Perception and Human Performance: Cognitive Processes and Performance,* vol.2, (pp. 43.1-43.39). New York: Wiley.

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced complacency. *The International Journal of Aviation Psychology, 31(3),* 1-23.

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

Parush and Auerbach (2005) *Proceeding of the HCI International 2005 Conference*, Las Vegas, NA, July 22-27, 2005.

Pashler, H.E. (1998). *The psychology of attention.* Cambridge, MA: MIT Press.

Prinzel, L.J., Freeman, F.G. , Scerbo, M.W., Mikula, P.J., & Pope, A.T. (2003). Effects of a psychophysiological system for adaptive automation on performance, workload, and the event-related potential P300 component. *Human Factors*, *45*(4), 601-613.

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

Rouse, W.B. (1991). *Design for Success,* New York: Wiley.

Sarter N.B. & Woods, D.D. (1994). Pilot interaction with cockpit automation II: An experimental study of pilots' model and awareness of the flight management system. *International Journal of Aviation Psychology, 4*, 1-28.

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

Sarter, N.B. (2000). The need for multisensory interfaces in support of effective attention allocation in highly dynamic event-driven domains: The case of cockpit automation. *The International Journal of Aviation Psychology, 10*(3), 231-245.

Scallen, S.F. & Hancock, P.A. (2001). Implementing adaptive function allocation. *The International Journal of Aviation Psychology, 11(2),* 197-221.

Scerbo, M.W. (1996). Theoretical perspectives on automation. In R. Parasuraman and M. Mouloua (Eds.), *Automation and Human Performance,* (pp. 37-63). Mahwah, NJ: Lawrence Erlbaum.

Schmorrow, D. D. (2002). DARPA's augmented cognition program: Tomorrow's human computer interaction from vision to reality: Building cognitively aware computational systems. In *Proceedings of the IEEE 7th Human Factors Meeting,* Scottsdale, AZ.

Schweizer, K. (1996). Levels of encoding, preattentive processing, and working-memory capacity as sources of cognitive ability. *Personality and Individual Differences, 21*(5), 759-766.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. Journal of Experimental Psychology: General, 125, 4-27.

Sheridan, T. (1972). On how often the supervisor should sample. *IEEE Transactions on Systems, Science, and Cybernetics, SSC-6,* 140-145.

Sheridan, T.B. (1988). Task allocation and supervisory control. In M. Helander (Ed.), *Handbook of Human-Computer Interaction*. Amsterdam: North-Holland.

Sheridan, T. B. & Verplank, W. (1978). *Human and Computer Control of Undersea Teleoperators.* Cambridge, MA: Man-Machine Systems Laboratory, Department of Mechanical Engineering, MIT.

Simpson, C. & Williams, D.H. (1980). Response time effects of alerting tone and semantic context for synthesized voice cockpit warnings. *Human Factors, 22,* 319-330.

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. Cognitive Psychology, 33, 5-42.

Sorg, B. & Whitney, P. (1992). The effect of trait anxiety and situational stress on working memory capacity. *Journal of Research in Personality, 26,* 235-241.

Sorkin, R.D. (1987). Design of auditory and tactile displays. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 549-576). New York: Wiley.

Speilberger, C.D., Gorsuch, R.L., Lushene, R.E., Vagg, P.R., & Jacobs, G.A. (1983). *Manual for the state-trait anxiety inventory (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C.D. (1972). Anxiety as an emotional state. In , C.D. Spielberger (Ed.), *Anxiety: Current Trends in Theory and Research.* London: Academic Press.

Spence, C. & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 22*(4), 1005-1030.

St. John, M., Kobus, D.A., & Morrison, J.G. (2002). A multi-tasking environment for manipulating and measuring neural correlates of cognitive workload. In *Proceedings of the IEEE 7th Conference on Human Factors and Power Plants*, 7-10 – 7-14.

Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.

Stokes, A.F. & Kite, A. (1994). *Flight stress: Stress, fatigue, and performance in aviation.* Aldershot: Avebury Aviation.

Szalma, J.L., Warm, J.S., Matthews, G., Dember, W.N., Weiler, E.M., & Meier, A. (2004). Effects of sensory modality and task duration on performance, workload, and stress in sustained attention. *Human Factors, 46*(2), 219-233.

Thropp, J.E., Oron-Gilad, T., Szalma, J.L. & Hancock, P.A. Incorporating individual differences into adaptive automation. Manuscript submitted.

Tso, K.S., Tharp, G.K.. Tai, A.T., Draper, M.H., Calhoun, G.L., & Ruff, H.A. (2003). A human factors testbed for command and control of unmanned air vehicles. In *Proceedings of the 22nd Digital Avionics Systems Conference*, Indianapolis, IN.

Tuholski, S. W., Engle, R. W., & Baylis, G. C. (2001). Individual differences in working memory capacity and enumeration. *Memory & Cognition, 29*(3), 484-492.

Tulga, M.K. & Sheridan, T.B. (1980). Dynamic decisions and workload in multitask supervisory control. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-10,* 217-232.

Validation of a comprehensive stress state questionnaire: Towards a state "big three"? In I. Mervielde, I.J. Deary, F. DeFruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (vol. 7), (pp. 335-350).

Warm, J.S. & Jerison, H.J. (1984). The psychophysics of vigilance. In J.S. Warm (Ed.), *Sustained attention in human performance* (pp. 15-59), Chichester, UK: Wiley.

Warm, J.S. (1984). *Sustained Attention in Human Performance.* London: Wiley.

Warm, J.S., Dember, W.N., & Hancock, P.A. (1996). Vigilance and workload in automated systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 183-200). Hillsdale, NJ: Erlbaum.

Weiner, B. & Schneider, K. (1971). Drive versus cognitive theory: A reply to Boor and Harmon. *Journal of Personality and Social Psychology, 18*, 258-262.

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

Wickens, C.D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and Performance VIII* (pp. 239-257). Hillsdale, NJ: Erlbaum.

Wickens, C.D. (1992). *Engineering Psychology and Human Performance,* New York: Wiley.

Wickens, C.D. (1999). The tradeoff in the design for routine and unexpected performance: Implications for situation awareness. In M.R. Endsley & D.J.

Garland (Eds.), *Situation awareness analysis and measurement.* Mahwah, NJ: Erlbaum.

Wiener, E.L. & Curry, R.E. (1980). Flight deck automation: Promises and problems. *Ergonomics, 23*(1), 995-1011.

Wiener, E.L. (1989). Human factors of advanced technology (“glass cockpit”) transport aircraft. (NASA Contractor Rep. No. 177528), Ames Research Center, Moffett Field, CA.

Williges, R.C., Williges, B.H., & Fainter, R.G. (1988). Software interfaces for aviation systems. In E. Wiener & D. Nagel (Eds.), *Human Factors in Aviation* (pp. 463-493). San Diego, CA: Academic Press.

Woods, D.D. (1996). Decomposing automation: Apparent simplicity, real complexity. In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human Performance: Theory and* Applications (pp. 1-16). Mahwah, NJ: Erlbaum.

Yarbus, A.L. (1967). *Eye movements and vision.* New York: Plenum.

Yeh, Y. Y., & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors, 30*, 111-120.

Young, M.F. & McNeese, M.D. (1995). A situated cognition approach to problem solving. In P. A. Hancock, J. Flach, J. Caird, & K. Vicente (Eds.), *Local Applications of the Ecological Approach to Human-Machine Systems* (pp. 359-391). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.