2012

# Comparing Types Of Adaptive Automation Within A Multi-tasking Environment

Grant S. Taylor
*University of Central Florida*

# COMPARING TYPES OF ADAPTIVE AUTOMATION WITHIN A MULTI-TASKING ENVIRONMENT

by

GRANT S. TAYLOR
M.A. University of Central Florida, 2010
B.S. University of Central Florida, 2006

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Applied Experimental/Human Factors
in the Department of Psychology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2012

Major Professor: James L. Szalma

# ABSTRACT

Throughout the many years of research examining the various effects of automation on operator performance, stress, workload, etc., the focus has traditionally been on the level of automation, and the invocation methods used to alter it. The goal of the current study is to instead examine the utilization of various types of automation with the goal of better meeting the operator's cognitive needs, thus improving their performance, workload, and stress. The task, control of a simulated unmanned robotic system, is designed to specifically stress the operator's visual perception capabilities to a greater degree. Two types of automation are implemented to support the operator's performance of the task: an auditory beep aid intended to support visual perception resources, and a driving aid automating control of the vehicle's navigation, offloading physical action execution resources. Therefore, a comparison can be made between types of automation intended to specifically support the mental dimension that is under the greatest demand (the auditory beep) against those that do not (the driving automation). An additional evaluation is made to determine the benefit of adaptively adjusting the level of each type of automation based on the current level of task demand, as well as the influence of individual differences in personality.

Results indicate that the use of the auditory beep aid does improve performance, but also increases Temporal Demand and Effort. Use of driving automation appears to disengage the operator from the task, eliciting a vigilance response. Adaptively altering the level of automation to meet task demands has a mixed effect on performance and workload (reducing both) when the auditory beep automation is used. However, adaptive driving automation is clearly detrimental, causing an increase in workload while decreasing performance. Higher levels of Neuroticism are related to poorer threat detection performance, but personality differences show no indication of moderating the effects of either of the experimental manipulations. The results of this study show that the type of automation implemented within an environment has a considerable impact on the operator, in terms of performance as well as cognitive/emotional state.

Dedicated to my parents, who provided the nature and nurture that made me who I am,

and to Kim, my wife, for helping me discover who I want to become.

# ACKNOWLEDGMENTS

This work was supported in part by the U.S. Army Research Laboratory (ARL) (W91CRB08D0015). The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of ARL or the U.S. Government.

I would like to thank the members of my committee: Dr. Lauren Reinerman-Jones, Dr. Peter Hancock, Dr. Mustapha Mouloua, and especially Dr. James Szalma for serving as my academic advisor throughout my graduate career. I would also like to thank the many members of IST's ACTIVE Laboratory who offered their assistance with this project, with special thanks to Dr. Stephanie Lackey, Daniel Barber, Kimberly Sprouse, and Brandon Sollins. Finally, thank you to the many friends and family members who have offered your support throughout this challenging process, your enthusiastic encouragement is greatly appreciated.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# INTRODUCTION

Modern Warfighters use and rely on increasingly complex systems to support their missions. The advanced functionality of these systems inherently results in greater complexity, but this functionality brings with it a greater need for system designers to strive for an appropriate match between the functionality of these systems and the operator's needs and abilities. The complexity of modern technology can easily overwhelm an operator (Cummings & Guerlain, 2007), resulting in an overall decrease in system effectiveness rather than the desired increase. This decline in system effectiveness can result in injury or even death for the fielded Warfighter.

More than 6,000 unmanned ground vehicles have been deployed in military operations in Iraq and Afghanistan (Pitts, 2009), and their numbers are expected to grow exponentially in the near future (U.S. Army UAS Center of Excellence, 2010). The growing popularity of unmanned systems is evidence of their many benefits – acting as force multipliers, extending manned capabilities, and allowing Soldiers to conduct their missions from relative safety (Barnes, Parasuraman, & Cosenzo, 2006). These tele-operational tasks are likely to become more prevalent as unmanned system capabilities increase and implementation costs decrease, resulting in a greater proportion of our military action depending upon these remote operator control interfaces. The operators of these unmanned systems can be placed in control of several vehicles simultaneously (Liu, Wasson, & Vincenzi, 2008; Saqer, Visser, Emfield, Shaw, & Parasuraman, 2011; Squire, Trafton, & Parasuraman, 2006), or given responsibility for secondary tasks concurrent with their vehicle control task, which risks pushing their cognitive faculties to or beyond their limits (Cummings & Guerlain, 2007). Essentially, the complexity of these unmanned systems are capable of generating an information stream that can quickly become more than the operator can handle. One simple method for reducing the cognitive load on the operator, while maintaining system efficiency, is to automate certain task components.

## Automation

Automation, as defined by Parasuraman and Riley (1997, p. 231), is "the execution by a machine agent (usually a computer) of a function that was previously carried out by a human." The widespread implementation of modern automation did not begin until the Industrial Revolution. In this period, technological advancements were developed specifically for the sake of automating tasks that were once complex human-operated responsibilities. For example, steam-powered engines allowed complex mechanical tasks to be performed with virtually no human intervention, short of continually fueling the engine. These advancements in automation technologies were one of the primary factors responsible for the advances in economic prosperity over the past two centuries, allowing for the creation of manufactured goods with far greater efficiency than when humans were required to perform the work. The resulting decrease in production costs allowed the goods to be sold at lower prices, drastically improving the economic standing of society as a whole.

A similar revolution was also led by the development of modern computers. However, the automation of this era is "smarter" than that of the Industrial Revolution. Previously, mechanical devices replaced humans to perform physical labor. In our modern age, electronic computers are automating cognitive tasks. Initially, these were relatively simple, such as math problems or other highly-structured, logic-based functions (Campbell-Kelly & Aspray, 2004). Recently, computer technology has sufficiently advanced to become capable of automating tasks in more complex domains such as aviation (Amalberti, 1999), driving (Stanton & Young, 1998), manufacturing (Groover, 2007), and medicine (Thompson, 1994).

Automation in complex systems provides accuracy and speed advantages that cannot be achieved by a human. Therefore, automating all possible tasks would seem to lead to the best system with optimum performance and efficiency. However, research has repeatedly shown that this is not the case (Parasuraman, 1987; Sheridan, 1997). The use of automation can result in unexpected negative outcomes including a loss of efficiency, performance, and safety, and thus a thorough understanding of precisely how system characteristics influence the operator (and vice-versa) is necessary before automation can be implemented appropriately.

**Problems with Automation**

The common thread underlying all problems with automation is that the automated process will never be perfectly reliable. Even relatively simple automated tasks will inevitably experience a failure (Parasuraman, 1987) as automated processes are nested within a larger system vulnerable to external influences. Therefore, the human operator needs to retain a central role in the system, as they must monitor and interact with the automation. This human-automation relationship should be carefully maintained to ensure optimal system performance. Particular concerns are misuse, disuse, and abuse of automation (Parasuraman & Riley, 1997), delegation of the human operator to supervisory control (Sheridan, 1992, 1997), and skill degradation (Mitchell, Cummings, & Sheridan, 2004).

### Misuse, Disuse, and Abuse

First described by Parasuraman and Riley (1997), misuse, disuse, and abuse refer to the various suboptimal ways in which automation is utilized, both by operators and system designers. Misuse refers to the operators' overreliance on an automated system. The operators develop an unrealistically high level of trust in the automation, and, as a result, fail to critically monitor its actions for potential failures (Parasuraman & Riley, 1997). For example, Riley (1994) found that nearly half of airline pilots failed to discontinue the use of an automated system after it committed an error that negatively impacted the performance of their task. This trend is not restricted to the laboratory (Young & Stanton, 2001), as incident reports have determined numerous fatal accidents resulting from an operator's negligence to recognize failures in automated flight systems (NTSB, 1973; Mouloua, Gilson, & Koonce, 1997; Lee & See, 2004). Mosier, Skitka, and Korte (1994) report that crew complacency, resulting from overreliance on the automated flight systems, is a factor identified in 77% of flight incident reports.

Disuse is essentially the opposite problem of misuse, and it occurs when operators underestimate the reliability of automation and therefore distrust it (Parasuraman & Riley, 1997). Disuse can result in the implementation of automation causing an increase, rather than the desired decrease, in operator workload. Under these circumstances operators may choose to manually perform the now-automated task components in order to continually monitor the distrusted automation for errors (Bainbridge, 1983). Disuse is particularly problematic in systems with automated alerts to warn of events which have a relatively low probability of occurrence, but a very high cost when they do. For example, complete engine failure in an airplane is a very unlikely event, but when it does occur it can result in great costs from the loss of both life and equipment (Netherlands Aviation Safety Board, 1992). Given the cost associated with this event, any automated system designed to monitor for engine failure will do so with a very lenient decision criteria, accepting a high number of false alarms in order to avoid a very costly miss (Poor, 1994). This lenient decision criteria is selected because the high number of false alarms are considered to have virtually no associated cost, short of the time needed for the crew to determine the cause for the alarm and disable it. However, repeated false alarms do in fact come with a greater expense. That cost is the potential disuse of a system because the operator's trust is negatively affected (Parasuraman, Hancock, & Olofinboba, 1997), demonstrated by the "cry wolf effect" (Bliss, 1997). This disuse can lead to the operator ignoring future alarms, negating the purpose of the system and resulting in an increased risk probability.

The problems with automation extend beyond the realm of the operator to the decisions made by system designers regarding the implementation of automation, leading to possible abuse (Parasuraman & Riley, 1997). Typically, automation is implemented in any task where there is an anticipated financial gain, through an increase in either production rate or accuracy of performance over that of a human operator. However, without diligent consideration of the impact this automation will have on the human operator, its introduction could decrease overall system performance. For example, if the specific functions of the automation are not known by the operator, they may perform actions that are incongruous or contrary to those of the automation, resulting in, at best, inefficiency, and at worst, a catastrophic system failure (Riley, 1996).

### Supervisory Control

Another consideration is the change in control experienced by the operator. In automated systems, the human operator who formerly performed the now-automated task is typically delegated to a supervisory control position (Sheridan, 1997, 2002). Therefore, rather than removing the task from the operator's responsibility, it simply changes the type of work the operator must perform (Edwards, 1977; Parasuraman & Riley, 1997). Specifically, the operator's role becomes one of monitoring the automation for errors, and if/when the automation fails the responsibility of performing the task is returned to the operator. However, because the operator has been taken "out-of-the-loop," and no longer maintains direct control over the task on a regular basis, their ability to perform is likely to degrade (Kaber, Omal, & Endsley, 1999; Mitchell, Cummings, & Sheridan, 2004). This is an increasingly common problem due to the increase of human operators delegated to supervisory control tasks as automation has become more prevalent (Lee & Moray, 1994). This shift, described by Hopkin (1992), is a change from tactical to strategic orientation, placing the operator in what is essentially a vigilance task (Noyes, 2009).

### Vigilance

Vigilance, or sustained attention, is the ability of an observer to detect and respond to infrequent critical signals amidst non-critical events over an extended period of time (Davies & Parasuraman, 1982; Reinerman-Jones, Matthews, Langheim, & Warm, 2011). Research has consistently demonstrated that humans are particularly poor vigilance performers (Mackworth, 1948; See, Howe, Warm, & Dember, 1995; Szalma et al., 2004; Warm, Dember, & Hancock, 1996). This results from the counterintuitive concept that vigilance is a capacity-draining, mentally demanding task (Parasuraman, Warm, & Dember, 1987; Warm et al., 1996). In other words, most vigilance tasks require the operator to monitor sensory information, deciding whether a signal is present or absent, leading to a progressive decline in performance (Parasuraman, 1986). The introduction of an automated system does add processing load to the operator because the task now requires an understanding of the automation's functions in addition to the components already required for task completion. However, there are instances in which cognitively demanding tasks do not elicit a loss of performance; rather performance stays the same or can even improve (See et al., 1995). Automated systems often fall in this latter category as they usually require operators to complete multiple tasks simultaneously, which are typically cognitive in nature.

**Implementation of Automation**

Despite the aforementioned concerns presented, the history of automation has shown it can be effectively implemented. The goals for automation support its utility and can be summarized into two primary categories: reducing cost (both financial and labor hours) and improving performance of the system (Parasuraman, 1987; Wiener, 1984; 1985).

While neither human nor automation can ever achieve perfectly efficient and accurate performance, when implemented appropriately they can work synergistically to achieve a level of performance greater than either are capable of individually (Hancock & Parasuraman, 1992; Hancock, Parasuraman, & Byrne, 1996). To accomplish this, it is critical that system designers take care to assign functions to both the human operator and the automation for which they are best suited. Fitts (1951) published his seminal function allocation list more than a half-century ago (Table 1), which has since been refined by many experiments, informing system designers of best practices for distributing task load. Toward that end, more recent discussions have questioned the applicability of the original Fitts' list in modern systems (Hancock & Scallen, 1996; Woods, 2002) and whether such a strictly defined distribution can be appropriate across the variety of technologies available today (Dearden, Harrison, & Wright, 2000; Sheridan, 2000).

**Table 1.** Distribution of tasks best suited for humans and machines. Adapted from Fitts (1951).

| Humans surpass machines in the: | Machines surpass humans in the: |
|---|---|
| • Ability to detect small amounts of visual or acoustic energy<br>• Ability to perceive patterns of light or sound<br>• Ability to improvise and use flexible procedures<br>• Ability to store very large amounts of information for long periods and to recall relevant facts at the appropriate time<br>• Ability to reason inductively<br>• Ability to exercise judgment | • Ability to respond quickly to control signals, and to apply great force smoothly and precisely<br>• Ability to perform repetitive, routine tasks<br>• Ability to store information briefly and then to erase it completely<br>• Ability to reason deductively, including computational ability<br>• Ability to handle highly complex operations, i.e., to do many different things at once |

Successful function allocation can occur only after careful consideration of the balance between human and system control, and an important step in that process is identifying the appropriate cognitive component to support. Automation is typically organized into classifications based on the function of human information processing that it serves to support. The most widely accepted classification system represents human information processing as a four-stage model (Parasuraman, 2000; Parasuraman, Sheridan, & Wickens, 2000): sensory processing, perception/working memory, decision making, and response selection (Figure 1). Type of automation is classified based on the components of information processing it serves to support: information acquisition, information analysis, decision and action selection, or action implementation.

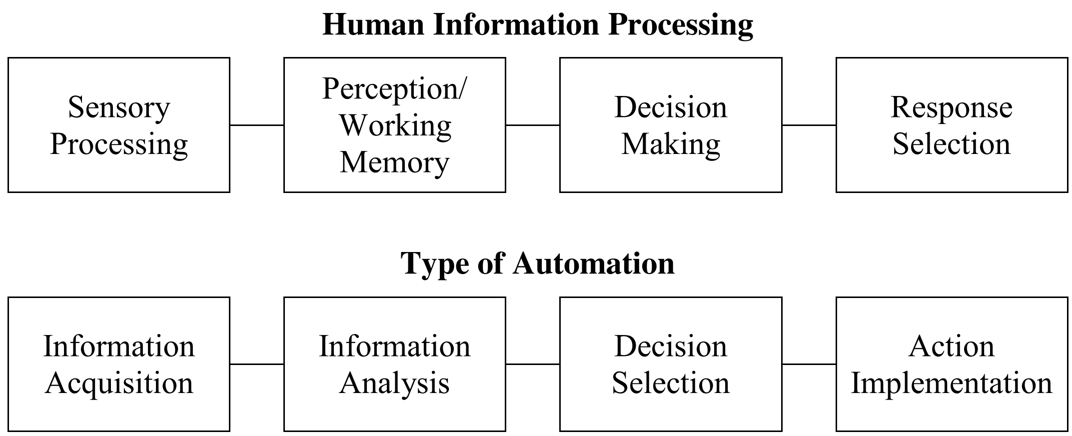

**Figure 1.** Simple four-stage model of human information processing and types of automation. Adapted from Parasuraman, Sheridan, and Wickens (2000).

This taxonomy consists of discrete categories of information processing stages and components; however, automation can support several aspects of information processing and exist at different levels for each of these functions. The level of automation describes the extent to which the automation takes control of the given function. A ten-level scale to describe the degree of autonomy of the system has received general acceptance (Parasuraman et al., 2000; Sheridan & Verplank, 1978), wherein higher levels represent greater automated control of the given task function (Table 2).

**Table 2.** Levels of automation of decision and action selection. Adapted from Parasuraman, Sheridan, and Wickens (2000).

| | | |
|---|---|---|
| | | The computer: |
| HIGH | 10 | 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 |
| | 7 | executes automatically, then necessarily informs the human |
| | 6 | allows the human a restricted time to veto before automatic execution |
| | 5 | executes that suggestion if the human approves |
| | 4 | suggests one alternative |
| | 3 | narrows the selection down to a few |
| | 2 | offers a complete set of decision/action alternatives |
| LOW | 1 | offers no assistance: human must take all decisions and actions |

## Adaptive Automation

In most traditional automated systems the type and level of automation are fixed at levels determined by the system designers. Some systems, such as an autopilot in an aircraft, allow the operator to activate or deactivate the automation at any time, providing *adaptable* automation (Opperman, 1994; Scerbo, 2001). *Adaptive* automation takes this concept a step further by dynamically altering the level of automation automatically (Hancock & Chignell, 1987). In this way, adaptive automation can be thought of as an additional layer of automation which encompasses the original system, serving to automate the operator's decision to turn the original automation on or off.

One goal of implementing adaptive automation is to avoid the problems inherent with static automation (misuse, disuse, skill degradation, etc.) while still reaping its benefits. As the term suggests, adaptive automation allows the automated aid to be adjusted responsively to better meet the needs of the system, including the human operator (Rouse, 1988). Through this method, the automation can be kept at a relatively low level during periods of routine performance, allowing the operator to maintain control without risking a reduction in overall system performance. However, when some aspect of the task changes that increases the demand on the human operator, or requires their attention to be devoted entirely to one specific sub-task, the system will respond by increasing the level of automation in certain areas. The goal for this adaptation is to effectively off-load some of the demands on the operator, allowing them to focus on critical elements.

The use of adaptive automation has been shown to be successful across various task environments. For example, Parasuraman, Cosenzo, and De Visser (2009) found the use of adaptive automation for a UAV control task to significantly increase operator situation awareness and change detection performance while reducing workload relative to both complete manual control as well as statically implemented automation. Adaptive automation has also been found to benefit the detection of system failures in a multitask flight simulation (Parasuraman, Mouloua, & Molloy, 1996; Parasuraman, Mouloua, Molloy, & Hilburn, 1993), an area which has specifically been shown to suffer from the implementation of static automation systems (Chambers & Nagel, 1985; Wiener, 1988). Beyond helping in the direct control of an aircraft, adaptive systems have also led to the reduction of workload for air traffic controllers (Hilburn, Jorna, Byrne, & Parasuraman, 1997; Kaber & Endsley, 2004). Although limited, adaptive automation has also exhibited benefits in other operational settings, such as the Rotorcraft Pilot's Associate used in Army helicopters (Dornheim, 1999; Miller & Hannen, 1999).

The ideal implementation of adaptive automation is for the system and operator to develop a relationship that more closely resembles a cooperative team rather than the system serving as the operator's tool (Hollnagel & Woods, 1999). For the system to be capable of adapting its level of automation to optimally support the operator's needs, it must maintain an accurate representation of the operator's cognitive state (Byrne & Parasuraman, 1996). It is for this reason that well-designed adaptive systems more closely resemble a coordinated team, in that effective team members understand their teammates' cognitive and affective state in order to adjust their own actions, better supporting the needs of others (Entin & Serfaty, 1999; Rouse, Cannon-Bowers, & Salas, 1992).

It is therefore necessary that the system maintain some representation of the operator's cognitive state so that it can appropriately determine when to adjust the level of automation (Wickens & Hollands, 2000). This can be accomplished through multiple means. The simplest method is to assume the operator's cognitive state based on external task conditions. For example, it can be assumed that a pilot's mental workload is highest during takeoff and landing, and so the level of automation in the cockpit can be increased during these times (Parasuraman, Mouloua, & Hilburn, 1999). However, this method has the obvious disadvantage of being unable to respond unexpected increases in workload, or differences across operators in their response to such changes.

An alternative is to infer the operator's needs based on their performance, with the assumption that the level of automation should be increased when performance begins to degrade (Kaber & Riley, 1999; Kaber, Wright, Prinzel, & Clamann, 2005). However, this method is difficult to implement as it requires the system to know the "true" state of the world in order to correctly evaluate the operator's performance. As this is typically impossible to achieve with a sufficient degree of accuracy, this method typically requires the implementation of a secondary task generated by the system (Kaber & Riley, 1999). The operator's performance on this secondary task can be monitored and used to infer the need for automation on their primary task. However, the implementation of this secondary task will serve as an additional source of cognitive demand and a distraction from the operator's primary task, making this method of adaptive automation control a costly one in terms of the operator's cognitive resources, and impractical in many operational settings. Given that the entire purpose of adaptive automation is to prevent an overloading of the operator's cognitive faculties, any implementation which inherently increases the cognitive demand of the task should be avoided if possible.

A third, and more promising, method used to infer the operator's cognitive state is the use of physiological measures. Measures of brain, heart, skin, and eye activity can all be used to estimate the operator's cognitive state (Kramer & Weber, 2000). This method has seen promise, with several researchers demonstrating its effectiveness in laboratory tasks (Bailey, Scerbo, Freeman, Mikulka, & Scott, 2006; Freeman, Mikulka, Pope, Prinzel, & Scerbo, 2003; Freeman, Mikulka, Prinzel, & Scerbo, 1999; Prinzel et al., 2003). However, this method is still not without its shortcomings. The primary difficulty resulting from the use of physiological measures is the amount of data necessary to make accurate predictions of cognitive state (Wickens & Hollands, 2000). These measures are incapable of providing truly real-time indications of the operator's mental state, as their calculations are always based on averages of data collected from some period of time. When averages are computed over longer time intervals, the resulting prediction of the operator's cognitive state may be more accurate, but this incurs the cost of introducing greater lag in the automation's response to the operator's needs.

## Operator Psychological Characteristics

### Stress

There is a long-recognized relationship between stress and performance. Traditional models define stress in terms of the relevant stimulus (e.g. noise, temperature, time pressure, etc.) or the physiological response elicited by this stimulus (Cox, 1978). Although the "inverted U" relationship between arousal and performance is often attributed to Yerkes and Dodson (1908), the first specification of this function as a description of arousal effects was by Hebb (1955; see Hancock & Szalma, 2003). Unitary arousal theory assumes that organisms seek to maintain an optimal level of physiological arousal, and that this level yields maximum performance (Figure 2).

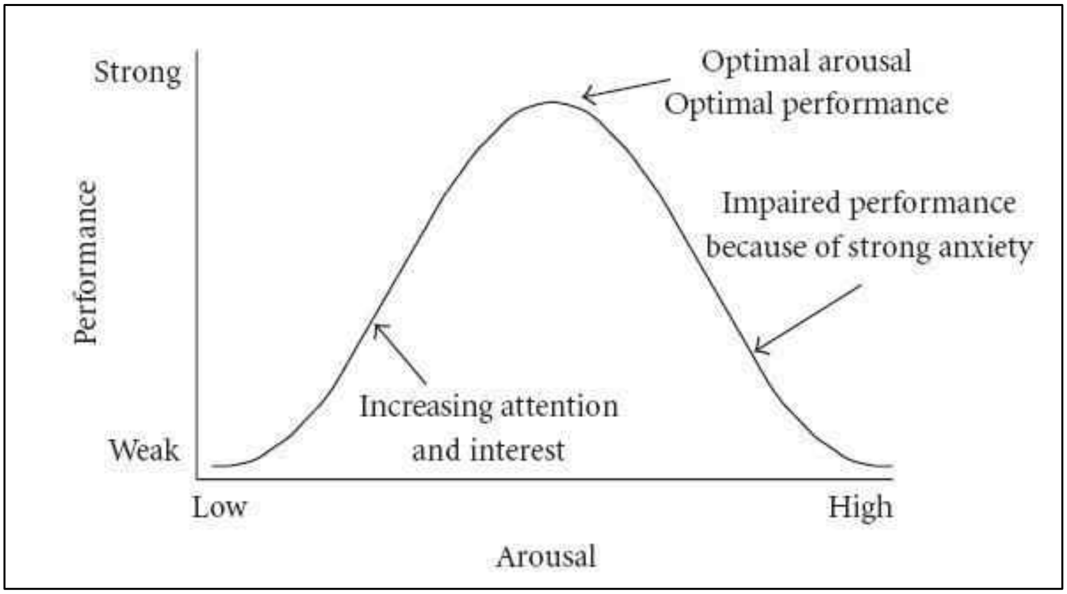


**Figure 2.** The "inverse U" relationship between arousal and performance. Figure from Hebb (1955).

More recent models, such as that proposed by Hancock and Warm (1989), have described a more complex relationship between stress and performance (Figure 3). The updated model maintains the central concept that a moderate level of stress is ideal to avoid performance decrements associated with under- or over-arousal. However, the model departs from the traditional view of the stress/performance relationship in two important ways.

First, the Hancock and Warm (1989) model recognizes that performance does not begin to degrade immediately when the operator's stress level is pushed above or below a narrowly defined ideal range. When stress levels are only marginally above or below the operator's "comfort zone", performance will (at least initially) be maintained due to psychological adaptability (i.e. increased attentional resources). If stress levels are pushed further, psychological adaptability will eventually fail due to the lack of additional attentional resources. At this point, physiological adaptability can provide for the continued maintenance of physiological functioning, until stress levels reach beyond the threshold of physiological adaptability. At this point, performance will begin to degrade dramatically as the operator no longer has the capacity to adapt to the continued hypo- or hyper-stress.

Secondly, the Hancock and Warm (1989) model incorporates the concept that stress is not simply a component of the environment which is imposed on to the operator. Rather, an interaction occurs between the task/environment and the operator to determine their response to its particular demands. Therefore, if the performer is hyperstressed and brought out of their comfort zone, the psychological adaptability which is necessary for them to maintain their performance will itself impose additional stress on the operator. For this reason, an apparently static level of external task demand can cause a consistently increasing, or decreasing, level of operator stress and ultimately result in a failure in performance, as is particularly evident in vigilance tasks (Warm et al., 1996). Of particular importance to the operator's response to task stress is the information rate (temporal flow) and information structure of the environment. Different individuals can perceive different meanings from tasks with identical information structures, resulting in different behaviors as well as different levels of stress.

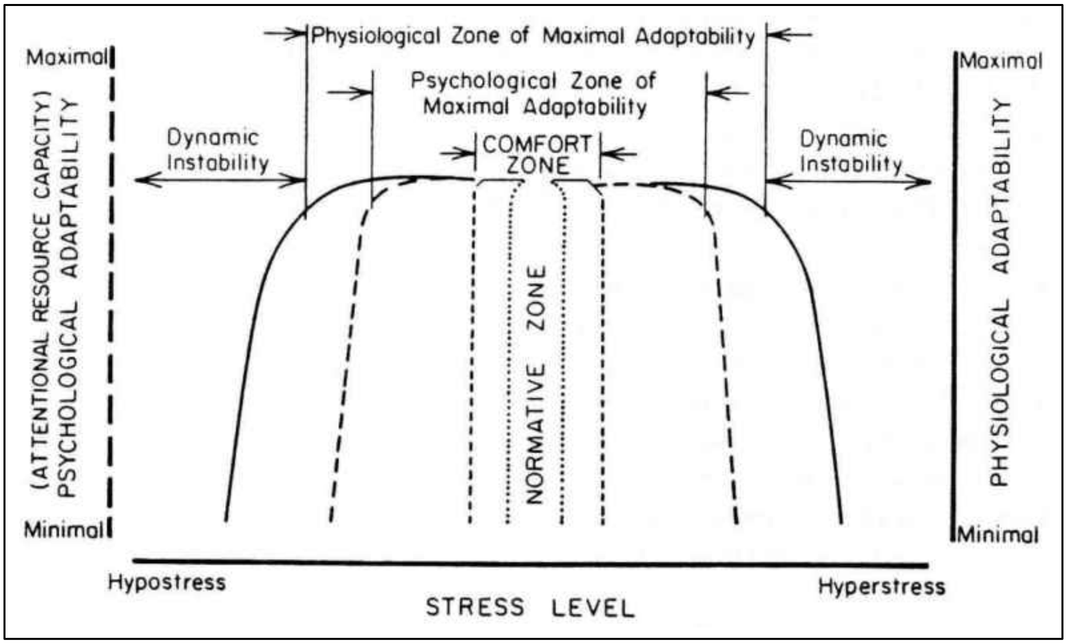


**Figure 3.** The Hancock and Warm (1989) model of stress as it relates to psychological and physiological adaptability.

### Multidimensionality of Stress

Matthews has further examined the relationship between performance and stress (Matthews et al., 2002, 1999). Rather than conceptualizing stress as a singular construct, Matthews developed a model consisting of three independent factors: Task Engagement (cognitive and energetic processes), Distress (cognitive and affective processes), and Worry (cognitive processes only; Matthews et al., 1999). These secondary factors are perhaps most easily understood through the primary factors of which they are comprised. Task Engagement consists of energy, motivation, and concentration; Distress consists of tension, hedonic tone, and confidence; and Worry consists of self-focus, self-esteem, task-relevant cognitive interference and task-irrelevant cognitive interference (Matthews et al., 1999). Each factor is associated with a core relational theme (Lazarus, 1991; Smith & Lazarus, 1990). Specifically, the Distress factor is related to the theme of perceived overload of processing capacity, Task Engagement is linked to the theme of commitment of effort, and Worry is related to the self-evaluation theme (Matthews et al., 2002).

By analyzing stress into three dimensions, a more thorough understanding of the relationship between stress and performance can be attained. Matthews and colleagues (2002) conducted a series of evaluations to examine how the performance of various tasks influences the three primary factors of stress. These findings indicated that single tasks often have different effects on each component (e.g. a visual vigilance task was found to decrease Engagement, increase Distress, and cause no significant change in Worry), and that the performance of different tasks result in different patterns of stress response (Figure 4). Task Engagement tends to be related to self-regulation, with tasks requiring high levels of short-term effort (working memory tasks) leading to increases in Engagement, while long, monotonous (vigilance) tasks led to decreases. Distress and Task Engagement were both related to the classic concept of arousal, which tended to increase over time, particularly in more demanding tasks. Worry consists of the self-evaluation and “meta-task” processes. In sustained attention, Worry tends to decrease over time as initial anxieties dissipate (see Szalma et al., 2004 for a more thorough discussion of the individual subscales of stress).

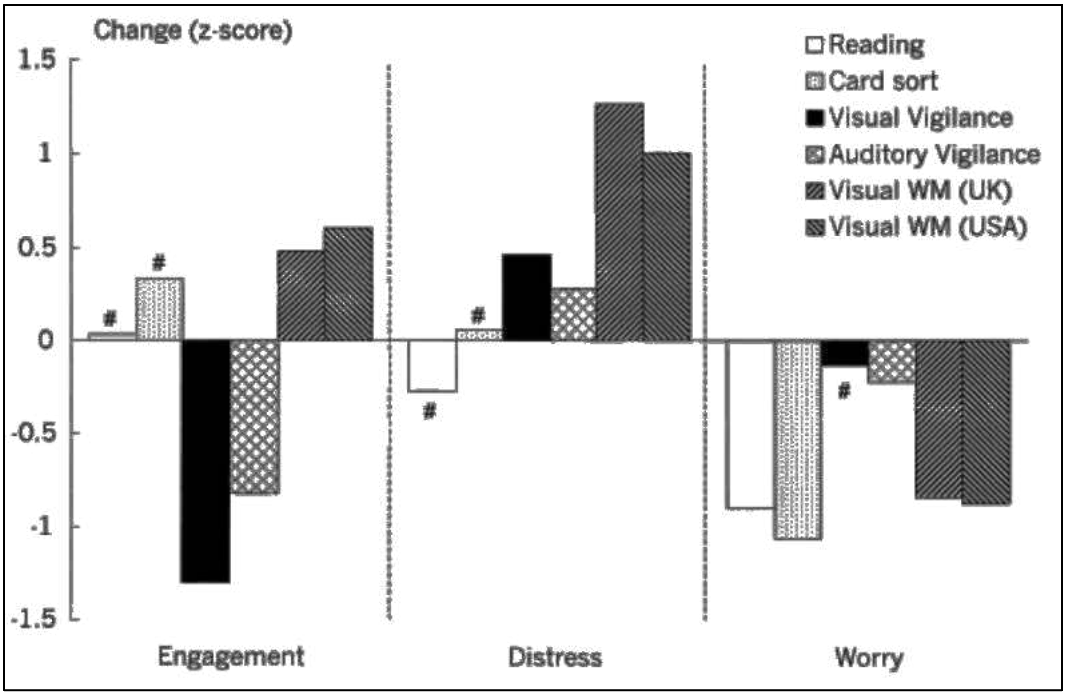


**Figure 4.** Task-induced changes in components of subjective stress resulting from various task types. Pound signs (#) indicate non-significant changes from baseline measures. Figure from Matthews et al. (2002).

### Application of Stress Theory to Adaptive Automation

The established relationship between stress and operator performance makes this concept an important consideration for adaptive automation systems. Static automation has been shown to effectively reduce Distress (Funke, Matthews, Warm, & Emo, 2007) and mitigate the decrease in Energetic Arousal (a component of Task Engagement) caused by vigilance tasks (Hitchcock et al., 2003), but only limited research has investigated the influence of adaptive automation on subjective stress (Szalma & Taylor, 2011). However, as the theoretical foundations of stress have evolved, it has become apparent that this relationship is not consistent across various task types, or even across different individuals completing a single task (Szalma, 2008). This complexity makes stress a more difficult construct to utilize as a means of controlling the level of automation. A stress-based adaptive automation system would inherently require the system to be specifically tailored to the particular tasks involved, as well as the state of individuals serving as operators. For these reasons, adaptive automation traditionally focuses on measures of operator workload rather than stress. Although stress may not be an ideal candidate to determine a system's level of automation, it is still a critical element to consider when evaluating the automated system's impact on the operator given its relationship with both performance and operator wellbeing.

## Workload and Resource Theories

Mental workload is the cognitive component most commonly used to trigger changes in the level of automation in adaptive systems. While there is no single, commonly accepted definition of workload, those that have been proposed conceptualize workload as the degree to which information processing, mental effort, or cognitive resources are required for task performance, relative to their capacity (Eggemeier, Wilson, Kramer, & Damos, 1991; Gopher & Donchin, 1986; Hockey, 1997; Kramer, Sirevaag, & Braune, 1987; Moray, 1979). Theoretical descriptions are usually metaphorical, typically invoking comparison to a hydraulic system in which tasks consume the fluid (resources) stored in a tank, or in economic terms wherein the cognitive resources are a limited commodity, subject to the demands of the current task(s) (Szalma & Hancock, 2007). Early perspectives considered cognitive resources as a single pool of energetic capacities (Kahneman, 1973; Figure 5). In contrast, others have argued for multiple resource capacities (Wickens, 1980, 1984).

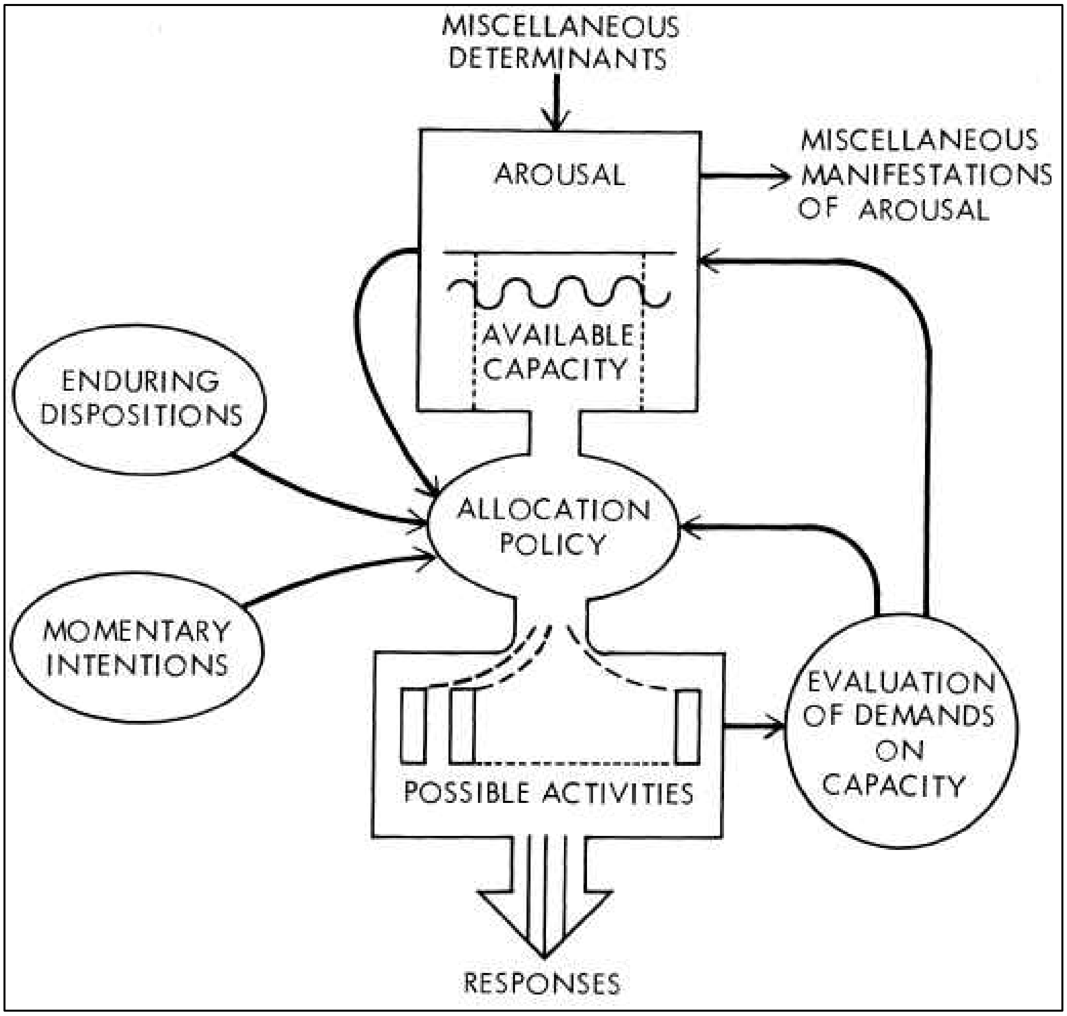


**Figure 5.** Kahneman's (1973) unitary resource theory.

Results of dual-task studies indicated that, for specific types of tasks, little to no detriment was caused by the introduction of a concurrent secondary task (Kantowitz & Knight, 1976; Wickens, 1976). For example, Wickens (1976) reported that the performance of a physical task was met with a degradation in the performance of a simultaneous manual tracking task (indicating that both rely on similar cognitive resources), while the performance of an auditory signal detection task caused no such degradation on the same manual tracking task. These results supported the multiple resource perspective that separate, unique pools of cognitive resources were responsible for the performance of the auditory task and the manual tracking task, allowing for their simultaneous performance at levels similar to that possible when performed individually.

### The Four Dimension Model

The model proposed by Wickens continues to be the most commonly accepted multiple resource theory, though Wickens himself conceded that it is not without its flaws (Wickens, 2008; see also Hancock, Oron-Gilad, & Szalma, 2007). Wickens' model is described as the Four Dimension (4-D) model, because it describes cognitive resources along four separate dimensions: stages of processing, codes of processing, modalities, and visual channels (Figure 6).

The stages of processing dimension divide tasks into perceptual, cognitive, and response phases. The codes of processing dimension separates tasks requiring spatial skills from those relying on verbal processes (both the perception and generation of speech). The modality dimension distinguishes visual from auditory perception, and for this reason is nested within the perceptual stage of processing, as it is not relevant for central processing or the selection and execution of actions. Finally, the visual channels dimension is nested within the visual modality, separating tasks dependent on focal vision (e.g. object recognition, reading text or symbols) from ambient vision (e.g. general orientation). These dimensions describe the separation of cognitive resources, and thus the extent to which two (or more) tasks can be completed simultaneously without sacrificing speed or accuracy. For example, two tasks which both utilize perceptual, spatial, visual, focal resources will strongly conflict with one another, but would cause virtually no interference with the verbal execution of actions.

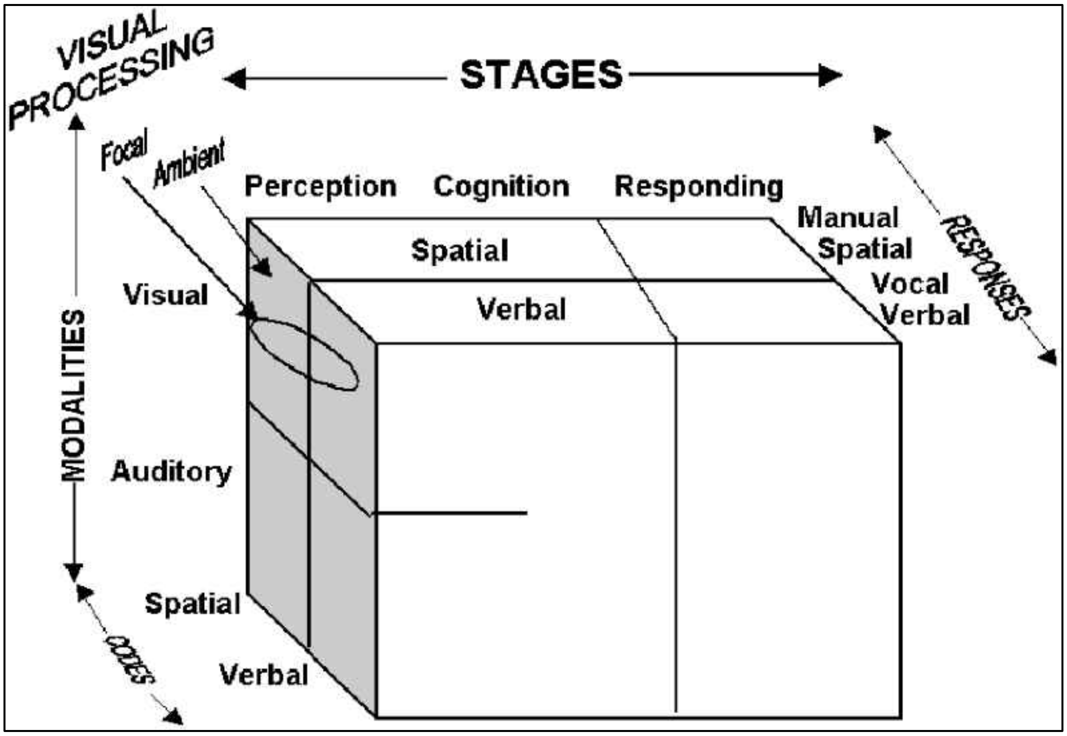


**Figure 6.** The 4-D multiple resource model. Figure from Wickens (2008).

**Individual Differences**

Although research has traditionally focused on the effects of various system characteristics on operator performance, workload, and stress, more recent work has investigated the importance of the operator's own characteristics in influencing response to automation. The most common trait evaluated has been trust (Lee & See, 2004), which has been found capable of influencing the operator's perceptions of automated systems more so than the reliability of the automation itself (Merritt & Ilgen, 2008). More recent work has begun to investigate the role other personality characteristics play in the operator-system relationship.

Szalma and Taylor (2011), building on previous evidence of a link between the personality traits Extraversion and Neuroticism with task performance (Eysenck & Eysenck, 1985), found these traits to influence an operator's interaction with automated systems as well. Neuroticism is a person's typical level of emotional stability, or their tendency to experience anxiety, anger, sadness, or guilt (Costa & McCrae, 1992), as well as the Distress and Worry components of stress (Matthews et al., 1999). Higher levels of Neuroticism have also been shown to negatively influence a person's ability to respond to dynamic task environments (i.e. fluctuating levels of task demand; Cox-Fuenzalida, Swickert, & Hittner, 2004), making it of particular relevance for adaptive automation systems. Szalma and Taylor (2011) found operators with higher levels of Neuroticism performed worse on a threat detection task, and were less likely to agree with an automated aid's correct recommendation.

Extraversion is primarily an index of a person's preference for social interaction, but is also sensitive to their preference for excitement and stimulation, as well as their assertiveness and positive affect (Costa & McCrae, 1992). Individuals higher in extraversion tend to have greater working memory and resource capacities, and superior divided attention, but poorer ability to sustain attention over time (Matthews, Deary, & Whiteman, 2003; Matthews, Jones, & Chamberlain, 1992). However, Szalma and Taylor (2011) found no statistically significant relationship between Extraversion and performance in their adaptive automation task. However, those higher in Extraversion were found to report higher levels of Frustration in periods of lower task demand than during periods of higher demand, suggesting that these participants prefer the more stimulating environment experienced during periods of high demand.

Findings such as these demonstrate the importance of considering individual differences in personality within automated systems, as their influence can be of equal or greater significance than the characteristics of the automated system. For example, an ideal system could respond to an operator who is particularly high in Extraversion by maintaining a higher general level of task demand relative to that preferred by individuals lower in Extraversion.

## Purpose of the Current Study

Research investigating the effects of automation on operator performance, stress, workload, etc. have tended to focus on the level of automation and the invocation methods used to alter it (Wickens, Li, Santamaria, Sebok, & Sarter, 2010). Despite repeated discussion of the importance of appropriate function allocation to avoid automation abuse (Dearden et al., 2000; Fitts, 1951; Hancock & Scallen, 1996; Sheridan, 2000; Woods, 2002), laboratory research studies still tend to select the task component to automate somewhat arbitrarily. Traditionally, this could be excused given the somewhat limited assortment of tasks that could be automated with acceptable levels of performance in operational environments, constraining the options of researchers who wished to maintain external validity. However, relatively few tasks remain that are incapable of being automated with some degree of reliability as a result of the continuing development of advanced technologies. A greater understanding of the specific impact of various types of automation is now needed given this growth in capabilities.

The limited research which has evaluated the impact of varying types of automation has done so only based on the stage of information processing they support (Parasuraman, 2000; Parasuraman et al., 2000), with little consideration of specific task demands. For example, there is evidence that operators receive greater benefit from adaptive automation applied to the information acquisition and action implementation stages of information processing and it has been argued that these effects are consistent across task types (Kaber, Perry, Segall, McClernon,

& Prinzel, 2006; Kaber et al., 2005). However, this generalized interpretation overlooks the possibility that these types of automation provided the greatest benefit because the operators experienced greater demands within the cognitive dimensions supporting information acquisition and action implementation. This alternative explanation would suggest that the type of automation which provides the greatest benefit to the operator cannot be universally defined, and is instead task-specific based on the extent to which a given task consumes mental resources of varying dimensions.

**Research Goals**

The traditional view of operator cognitive state is that workload and stress are unidimensional, but more recent work has established that these constructs are multidimensional (Matthews et al., 2002, 1999; Szalma et al., 2004). An operator can easily experience tremendous cognitive load within one specific dimension and relatively little along another. In such an instance, the type of automation implemented to support the operator may be critical to its success. The automation must be capable of supporting the relevant cognitive dimension, as anything that does not support this specific aspect of the task would likely provide little or no benefit.

The goal of the current study is therefore to advance the scientific understanding of the interaction between human operators and adaptive automation systems by manipulating the type and level of automation in the context of changing levels of task demand. The use of multiple simultaneous tasks in a simulated unmanned robotic system control environment will ensure that participants experience demand across all potential mental dimensions (Rouse, 1977) within a task that accurately reflects the type of complex task environment experienced by our current and future Warfighters. By focusing a particularly high level of demand within one specific cognitive/perceptual dimension, a comparison can be made between types of automation which support this mental dimension (demand-type matched automation) against those that do not (demand-type mis-matched automation), with the goal of improving operator performance, workload, and stress. An additional evaluation will be made to determine the potential benefit of adaptively adjusting the level of each type of automation based on the level of task demand, as opposed to maintaining a consistently high level of automation.

## Hypotheses to be Tested

**H1.** The use of automation which specifically supports the cognitive dimension experiencing the greatest level of demand (demand-type matched automation) will result in improved levels of performance, as well as reduced workload and stress when compared to alternative types of automation (demand-type mis-matched). Specifically, demand-type matched automation is expected to result in significantly reduced levels of the Mental and Temporal Demand subscales of workload as well as the Distress facet of stress.

**H2.** Adaptive automation, in which the level of automation adapts according to the current level of demand, will result in improved levels of performance and stress when compared to automation maintained at a consistently high level. This effect is further predicted to be stronger when the adaptive automation is implemented with demand-type matched automation, as compared to demand-type mis-matched automation. Although no prior research exists on which to base specific predictions regarding the stress effects, making this relationship more exploratory in nature, it is predicted that the adaptive automation will increase levels of Task Engagement while reducing levels of Distress.

**H3.** Individual differences in personality will moderate the relationship between the adaptability of automation variable with the dependent variables of performance, workload, and stress. Specifically, those higher in Neuroticism will receive less benefit from the adaptive automation, given their poorer ability to adapt to dynamic work environments (Cox-Fuenzalida et al., 2004). Conversely, those higher in Extraversion will receive greater benefit from the adaptive automation given their preference for more stimulating environments (Szalma & Taylor, 2011), which will be diminished when the automation remains at a constant high level.

# MATERIALS AND METHODS

## Experimental Task

The experimental task simulated the operation of an unmanned ground vehicle (UGV) from a remote operator control station, utilizing the Mixed Initiate Experimental (MIX) testbed (Figure 7; Barber et al., 2008; Reinerman-Jones, Barber, Lackey, & Nicholson, 2010). The mission took place in a generic Middle Eastern town, using a terrain database of the Military Operations in Urban Terrain (MOUT) site in 29 Palms, California. The task was completed on a standard desktop computer with a 22" (16:10 aspect ratio) monitor with a joystick and mouse. The participant was responsible for completing three separate tasks simultaneously: driving the vehicle along a pre-specified path, monitoring a video feed for enemy threats, and monitoring a map display for changes in entity locations.

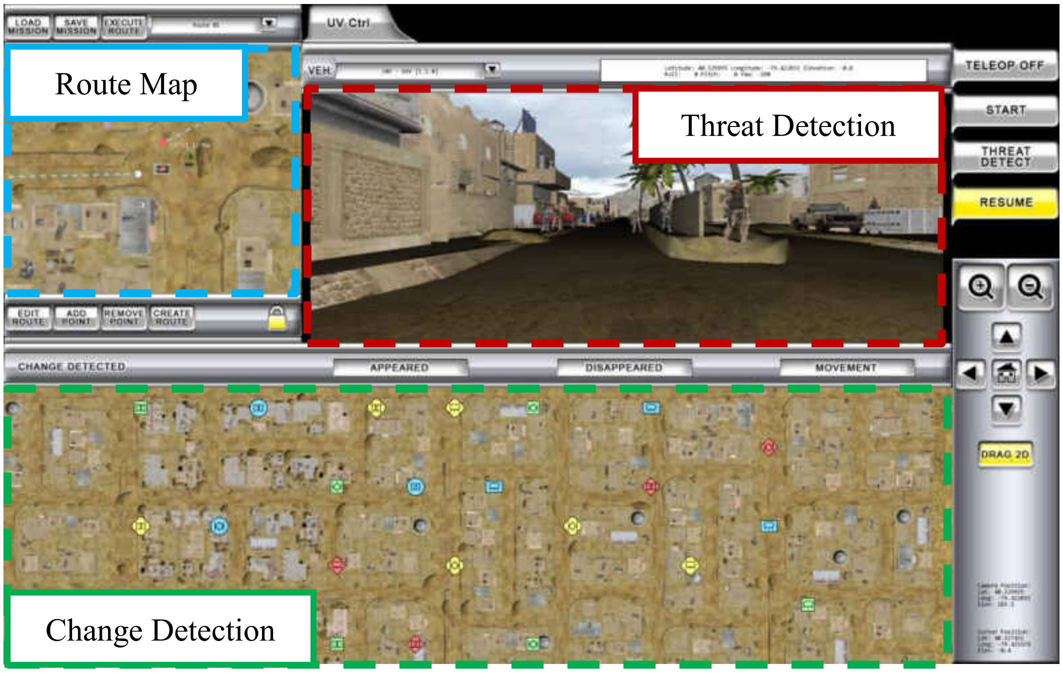

**Figure 7.** The MIX testbed, with outlines overlaid to differentiate task areas.

**Driving Task**

The participants' task was to follow a pre-defined path presented to them in the route map window (Figure 8). An icon representing the UGV's current location and heading was always displayed in the center of this window with North always at the top of the screen. Participants controlled the movement of the UGV using a joystick (Logitech Extreme 3D Pro) and the map continuously updated to follow the vehicle as it drove through the route. The route consisted of a series of waypoints (represented by large dots) connected by a dotted line.

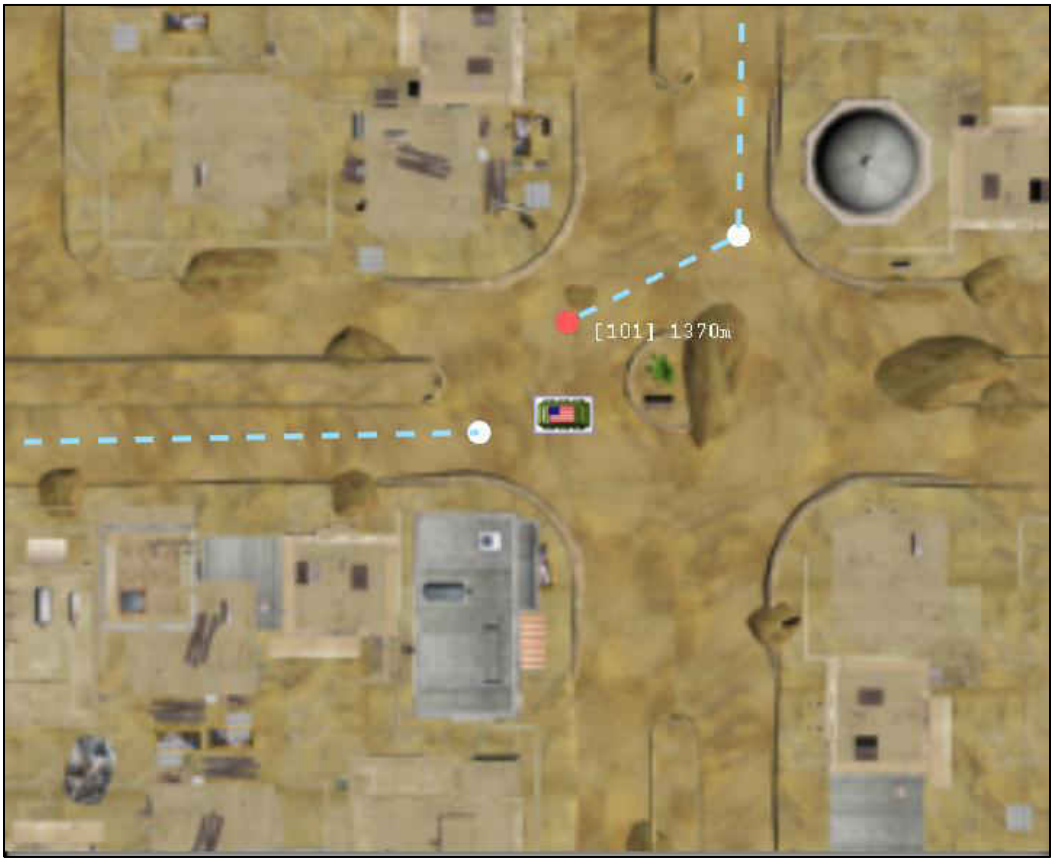

**Figure 8.** The route map portion of the MIX testbed, enlarged to show detail.

Four unique paths were used for the routes (Figure 9). There were two separate routes and each was used twice by reversing the start/end points. The routes had an equal number of turns (8 in each), with an equal number being left and right. Each route was 1.13 miles long and the UGV operated at a maximum speed of 2.82 mph, making each route last 24 minutes. This route design was implemented to balance and equate all features of the route paths in order to minimize the influence that the route could have on performance.

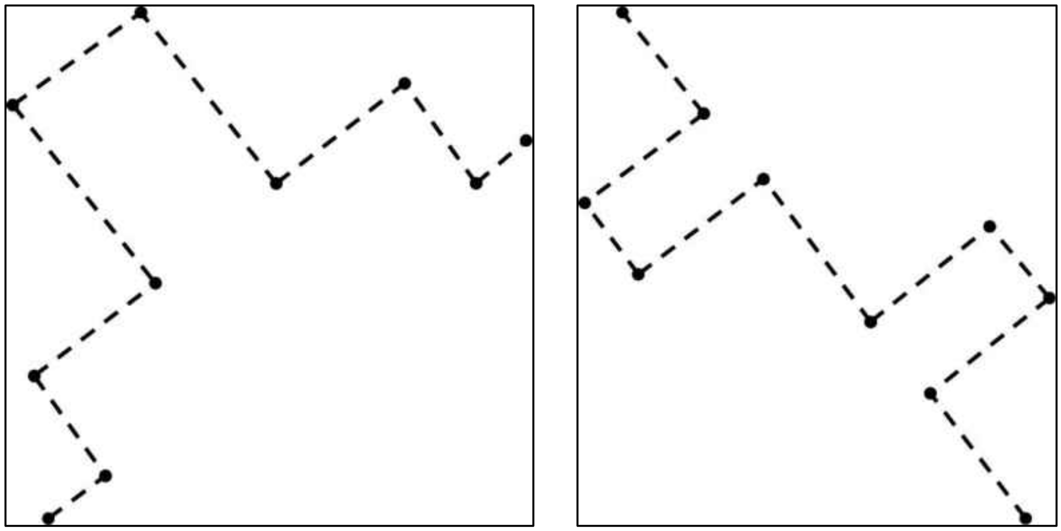

**Figure 9.** Routes used in the study. Each route is used with the start/end points reversed, creating a total of four unique paths.

Automation could be implemented in the driving task (see *Manipulations* section below). When driving automation was engaged, the UGV drove itself along the route automatically, with no input from the participant necessary. The participant did maintain a limited level of supervisory control through the use of a “Pause” button located on the right side of the MIX testbed control interface. Clicking this button with the mouse while the driving automation was engaged caused the UGV to stop in place until the participant clicked the same button again, which was labeled as “Resume” when the vehicle was paused. This type of control was similar to that of manual driving wherein the operator released the joystick to stop the vehicle and resumed driving by pushing the joystick forward.

### Threat Detection Task

As the vehicle drove along the route, a video feed from the perspective of the front of the UGV was displayed to the participant in the threat detection window (Figure 10). The environment was populated with various stationary objects, such as buildings, trees, vehicles, and people. The participant's task was to monitor the people along the route for potential threats. Four different categories of people were present in the environment (Figure 11): Friendly Soldiers, Friendly Civilians, Enemy Soldiers, and Insurgents (armed civilians). When an Enemy Soldier or Insurgent was visible, the participant was to identify them by clicking a button labeled "Threat Detect" and then clicking on the threat in the threat detection window using the mouse.

**Figure 10.** The threat detection portion of the MIX testbed, enlarged to show detail. A threat (Enemy Soldier) is visible on the right in green.

**Figure 11.** Examples of characters displayed throughout the environment. From left to right: Friendly Soldier, Friendly Civilian, Enemy Soldier, Insurgent.

The characters were presented to the participant at an average rate of 26 each minute, though this number could vary slightly dependent upon the speed with which the participant operated the UGV. Two of these 26 characters were classified as threats, resulting in a signal-to-noise ratio of 1:12. In addition to the human characters, neutral objects (e.g. rubble piles, vehicles, and trees) were also presented at an average rate of 15 per minute.

### Change Detection Task

A separate map at the bottom of the screen displayed the current location of various entities. Each entity was represented by an icon (Figure 12, Figure 13). Although these icons do convey information regarding its type and affiliation through military convention (Department of Defense, 2005), the participant was not trained or instructed to attend to these details. The participant's task was only to monitor the presence and location of the icons, and respond when a change occurred. Three types of changes occurred: Appear (a new icon was added to the display), Disappear (an icon was removed from the display), or Movement (an icon changed location). When one of these changes occurred, the participant responded by clicking the appropriate button ("Appeared," "Disappeared," or "Movement") above the change detection map.

Automation could also be implemented in the change detection task. When the automation was engaged, a beep alert sound was played over speakers at the moment any type of change occurred. The same beep sound was used regardless of the type of change that occurred. The beep was simply used as an alert and did not take control of the change detection task from the participant, as they were still required to respond whenever a change occurred.

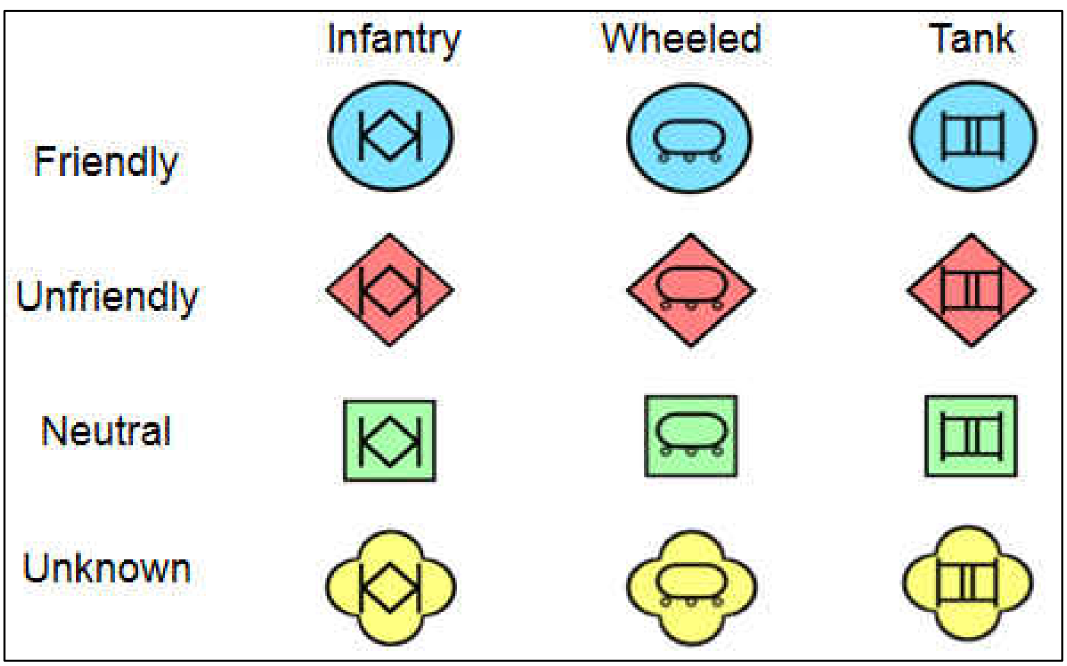


**Figure 12.** Icons used to represent entity positions for the change detection task.

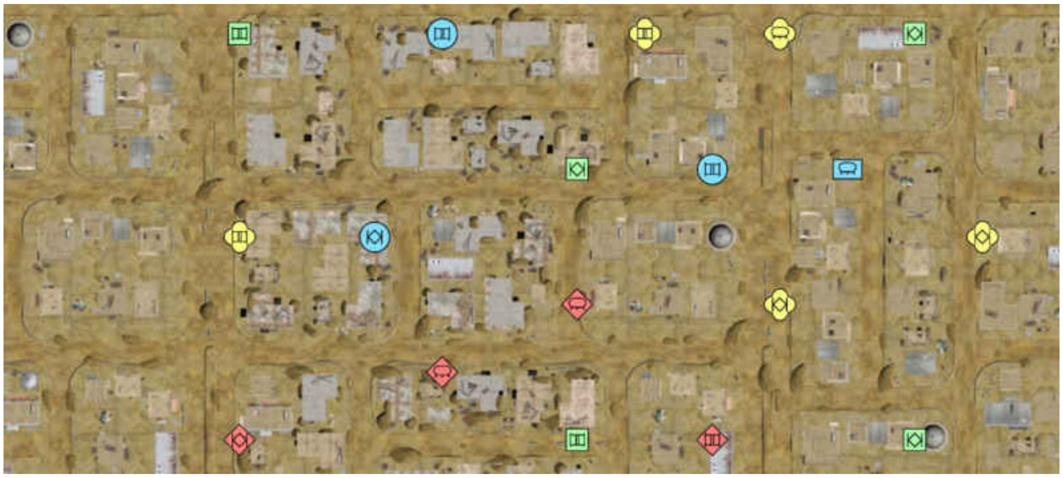

**Figure 13.** The change detection map, enlarged to show detail.

## Manipulations

### Task Demand

The level of task demand was manipulated at regular intervals between two levels: low and high. This manipulation altered the parameters of the change detection task in three ways: event rate, signal saliency, and working memory load. The event rate was manipulated between a slower rate (4 changes per minute) during low task demand and a faster rate (10 changes per minute) during high task demand. Both event rates are presented as averages, as the time elapsed between events varied within each condition to prevent the changes from occurring at expected intervals. Signal saliency was manipulated by adjusting the number of icons that change simultaneously whenever a change event occurred, with three icons changing at once during periods of low demand (with all three performing the same type of change – appear, disappear, or movement) and only one icon changing at a time for periods of high task demand. Working memory load was manipulated by adjusting the average number of icons present on the map at a single time. During periods of low task demand an average of eight icons were visible on the map at once, and during periods of high task demand this number was increased to an average of 24. The average number of icons is due to the continuous nature of the change detection task. To be clear, task demand condition was composed by the simultaneous occurrence of all three manipulations described above.

### Type of Automation

Participants were randomly assigned to receive one of the two types of automation: driving automation or beep alerts, as described in the *Experimental Task* section.

### Level of Automation

Each type of automation was implemented at two levels: low and high. When the automation was at a low level it provided no assistance, meaning the task dynamics were identical when the level of automation was low, regardless of the type of automation the participant was assigned. When the automation was at a high level, the assistance it provided varied as a function of the automation condition. Those participants in the driving automation condition had the vehicle drive itself along the route automatically, while those in the beep alert condition received auditory alerts whenever a change occurred on the change detection map.

### Static/Adaptive Automation

All participants experienced their assigned type of automation in both static and adaptive forms. When static, the level of automation maintained a consistently high level throughout a single experimental scenario. When adaptive, the level of automation fluctuated as a function of the level of task demand. However, the level of automation did not adjust immediately to a change in the level of task demand. In order to simulate the time required for an automated system to recognize that a change in the level of task demand had occurred, a consistent time delay elapsed after a change in the level of task demand before the level of automation changed to match it (see *Experimental Scenarios* section for details).

**Experimental Scenarios**

These manipulations combined to form four unique experimental scenarios, with each participant completing all four scenarios. Each of these scenarios lasted for 24 minutes, with changes to task parameters – task demand and level of automation (changes to the level of automation only occurred in adaptive scenarios) – occurring at three-minute intervals, dividing each scenario into eight blocks.

Each participant received two static automation and two adaptive automation scenarios. Two scenarios, one adaptive and one static, began under low task demand, with the remaining two scenarios beginning under high task demand. The level of task demand always alternated between low and high at 3, 9, 15, and 21 minutes. For example, scenarios starting under low demand changed to high demand at 3 minutes, then back to low demand at 9 minutes, high demand at 15 minutes, and return to low demand from 21 minutes through the end of the scenario. Scenarios that began under high demand followed the opposite pattern. Starting scenarios under both low and high demand allowed for any potential influence of the order of task demands to be negated.

In the two adaptive automation scenarios, automation began at a level matched to the initial level of task demand (low automation for low demand, high automation for high demand). The level of automation then adapted to the changing level of task demand throughout the scenario, but with a three minute delay. For example, if the level of task demand increased from low to high at the three minute mark, the automation maintained a low level until the six minute mark, at which point it increased to a high level. The level of task demand then returned to a low level at nine minutes, with the automation continuing at a high level until the 12 minute mark, when it decreased to a low level to match the level of demand (Figure 14).

The reason for this delay was to simulate the time needed for the system to detect a change in the operator's cognitive state. Although three minutes is slightly longer than physiological-based metrics of cognitive state typically require to detect changes in workload, the delay was intentionally over-estimated. Several studies have found evidence that adjusting the level of automation while an operator is performing a task can have a brief negative impact on their performance, workload, and situation awareness (Hilburn, Molloy, Wong, & Parasuraman, 1993; Kaber, Wright, & Sheik-Nainar, 2006; Parasuraman, Bahri, Molloy, & Singh, 1991; Reinerman-Jones, Taylor, Sprouse, Barber, & Hudson, 2011), and have therefore recommended that the level of automation not be adjusted immediately upon detecting a change in operator workload. Introducing a slight delay before changing the level of automation provides the system with adequate time to ensure that the newly-detected state will persist, reducing the risk of changing the level of automation (temporarily reducing operator performance) to meet a fleeting level of demand, only to soon return to the original level of automation (again temporarily reducing operator performance).

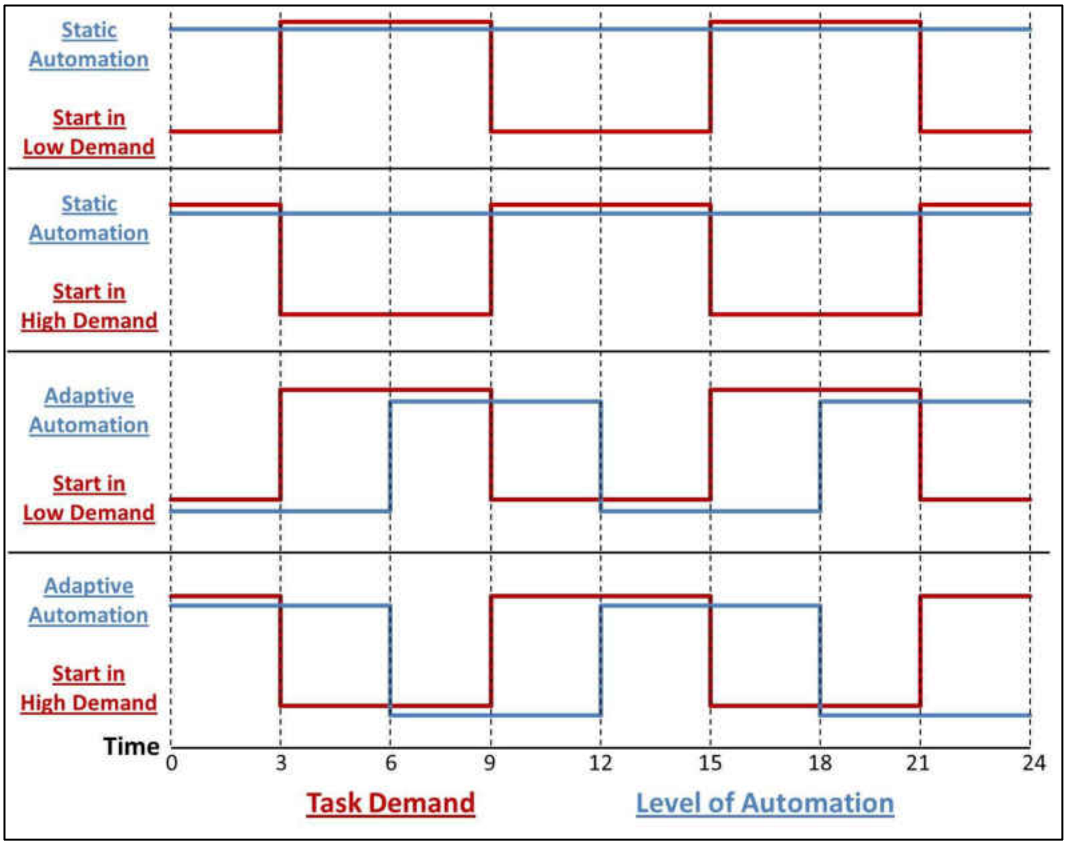


**Figure 14.** The changing levels of automation and task demand for the four experimental scenarios.

## Measures

### Questionnaires

#### Demographics

Participants completed a demographics questionnaire to measure standard items such as age and gender, as well as items used to determine their experience with various technologies. This questionnaire was also used to ensure that the participant met the inclusion criteria: normal state of health, normal color vision, and no prior military experience (APPENDIX A).

### Personality Measure

Items from the International Personality Item Pool (IPIP; Goldberg et al., 2006) were used to measure the participant's levels of Extraversion and Neuroticism, using 10 items for each trait presented in a fixed random order (APPENDIX B).

### Stress Measure

The Dundee Stress State Questionnaire (DSSQ; Matthews et al., 2002) was used to assess the participants' subjective stress levels following each experimental scenario. Due to time constraints, the short form was used, which produced measures of Task Engagement, Distress, and Worry. The DSSQ required a pre-test to be completed before beginning the experiment (APPENDIX C) and a post-test to be completed following each experimental scenario (APPENDIX D).

### Workload Measure

The NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988) was used to measure the participant's subjective workload from each experimental scenario. The measure produced six workload subscales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration Level, as well as a single combined measure of global Workload. The global Workload measure was calculated as the weighted average of the six subscales, with each subscale weighted according to the number of times it was selected as the more important contributor in the paired comparisons section. The NASA-TLX was administered on the computer through a standard computer program (APPENDIX E).

## Physiological Measures

### Electroencephalogram

A nine channel electroencephalogram (EEG) system from Advanced Brain Monitoring was used to record participant brain activity (Figure 15). The system sampled at 256 Hz from F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4 using the International 10-20 System, with references at each mastoid. Power spectral density analysis was used to compute values for Alpha (8-13 Hz), Beta (14-26 Hz), and Theta (4-7 Hz) activity at each sensor site.

Success has been found specifically within the adaptive automation domain through the use of a combined ratio of EEG activity to calculate engagement (Freeman et al., 1999; Pope, Bogart, & Bartolome, 1995). Rather than evaluating separate bands of EEG activity independently, activity from four sensor sites (Pz, Cz, P3, and P4) was combined across three common bands [beta / (alpha + theta)] to form a single value. This engagement index is based on previous findings that beta activity is sensitive to increases in arousal and attention, while alpha and theta can detect decreases (Abarbanel, 1999), and has been shown to be capable of successfully manipulating the level of automation within an adaptive system (Freeman et al., 2003). For this reason, the engagement index was calculated to evaluate the influence of the experimental task and manipulations on the participants' cognitive state.

**Figure 15.** The Advanced Brain Monitoring nine channel EEG system.

### Electrocardiogram

An electrocardiogram (ECG) system was used to measure participant heart activity. A Thought Technology ProComp Infiniti encoder was used with an ECG-Flex/Pro sensor that sampled at 2048 Hz. Three electrodes were attached to the participant's torso (Figure 16): one on both collar bones and one just below the sternum.

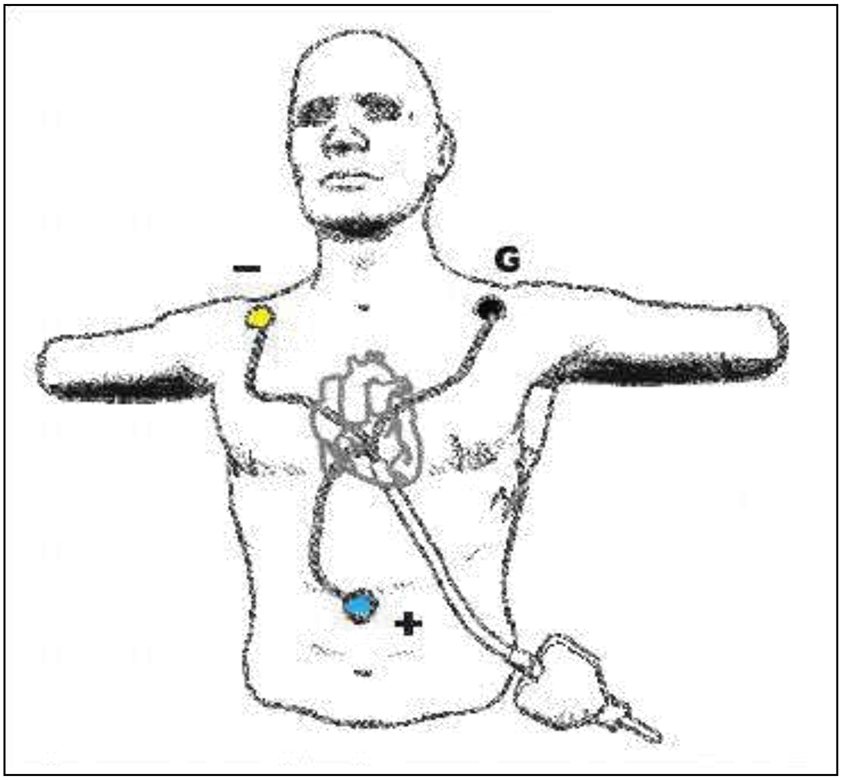


**Figure 16.** Sensor placement for the ECG system.

Two measures have traditionally been derived from ECG recordings to estimate cognitive workload: heart rate and heart rate variability (HRV), with heart rate increasing and HRV decreasing with increases in workload (Byrne & Parasuraman, 1996; Vicente, Thornton, & Moray, 1987; Wilson, 1992). Because measures of simple heart rate can be influenced by extraneous factors such as physical activity, it is considered to be less diagnostic than HRV. HRV was calculated as the statistical variance of the elapsed time of each heartbeat across a series of time, with decreases in variance indicating an increase in mental workload (Prinzel et al., 2003).

## Participants

University undergraduate students served as the experimental participants and were recruited using an experiment management website. The participants received credit for their psychology courses for completing the study. Given the sensitive nature of the terrain database used in the UGV control task, all participants were required to be US citizens. Participants were also required to be right handed (due to potential differences in brain physiology of left handed participants), have normal (or corrected to normal) vision, and have no prior military service. Participants were also asked not to consume alcohol or any sedative medication for 24 hours or caffeine for two hours prior to the study, as these could influence the physiological response recorded by the EEG and ECG.

## Experimental Procedure

Upon arrival, participants were first confirmed to be U.S. citizens through verification of their birth certificate, passport, or voter's registration card. After their citizenship was confirmed, the participant was provided with an Informed Consent form that detailed their rights as a research participant, the purpose of the study, an overview of the procedure, and the potential risks associated with participating.

The EEG cap was then placed on the participant. The cap was aligned using the nasion (the midpoint between the eyes, just above the bridge of the nose) and inion (the bump found at the center of the occipital bone on the back of the skull). If necessary, the participant's hair was parted at the site of each EEG sensor to ensure direct contact between the sensor and the scalp. Conductive gel was also used to ensure proper connection and to reduce the electrical impedance of the signal. In addition to the nine sensors, the system used two reference electrodes – one on each mastoid bone (behind the ear), which were attached directly to the participant's skin using adhesive pads. Once all sensors were in place, they were tested to confirm that the electrical resistance of each was below 30 kΩ.

The participant then completed the preliminary questionnaires: Demographics, DSSQ Pre-Test, and the Personality measure. Following these questionnaires, the researcher described the experimental task through a PowerPoint presentation. This training covered each portion of the task (driving, threat detection, and change detection) separately. Following each portion of the presentation, the participant completed a brief practice scenario in which they only performed one task at a time. After performing each task individually, the participant completed two full practice scenarios in which they performed all of the tasks simultaneously.

Following these practice missions, the ECG sensors were attached to the participant's collar bones and stomach using the same adhesive electrodes used for the EEG references. Once connected, the ECG signal was verified and the participant's baseline physiological activity was recorded for five minutes. The participant was asked to relax with their eyes open while the data was collected. The data recorded during this period was used as a baseline to which recordings made during the experimental scenarios were compared, accounting for the random variation in individual physiological differences.

After completing the resting baseline, the participant began their first full experimental scenario. The order in which all participants completed the experimental scenarios was counterbalanced using the *Latin Square* method to ensure that any influence the scenario order may have on performance, physiological response, or subjective workload/stress was evenly distributed across all four scenarios. Likewise, the routes used for each scenario were independently counterbalanced using a separate *Latin Square* design to ensure that any benefit or disadvantage that may come from any single route was also equally distributed across all four experimental scenarios.

After completing the experimental scenario the participant completed the DSSQ Post-Test measure and the NASA-TLX. This pattern repeated for the remaining three experimental scenarios. After the fourth experimental scenario and questionnaires, the EEG cap and ECG sensors were removed from the participant and they were allowed to leave. The entire experiment lasted two hours.

# RESULTS

## Power Analysis

Early pilot testing suggested that automation type had a fairly strong effect on performance, with an average effect size (Cohen's *d*) of $d \approx 1$. The adaptability of automation exhibited medium effect sizes, averaging $d \approx 0.5$. Based on these data, a power analysis determined that a total sample size of N = 70 (35 per automation condition) with $\alpha = .05$ would provide sufficient power for detecting between-subject effects of automation type as well as within-subject effects of automation adaptability ($1 - \beta = 0.984$ in each case).

## Sample Population

Data was collected from a total of 70 university undergraduates. However, errors associated with the simulation testbed as well as the physiological sensors required data from 10 participants to be removed. Although this reduction in sample size did reduce statistical power to reject the null hypotheses, the remaining sample still provided adequate power to detect effects of both automation type and adaptability (0.967 in each case). Of the 60 participants in the final dataset, there were 31 females (age: $M = 19.31$, $SD = 2.19$) and 29 males (age: $M = 19.78$, $SD = 3.47$). Of these, 29 participants received the driving automation, with the remaining 31 receiving the beep alert automation. A chi-square test confirmed that equivalent numbers of each gender were present in each experimental condition [$\chi^2(1, N = 60) = 0.258$, $p = .611$].

## Manipulation Check

An evaluation of the task demand manipulation was first conducted to confirm that this manipulation had the desired effect on change detection performance. Performance on the change detection task was measured through two values: percent of changes detected and percent of changes correctly identified. The percent of changes detected was calculated as the number of changes to which the participant made a response, regardless of whether they correctly classified the type of change that occurred (appear, disappear, or movement), divided by the total number of changes presented throughout the given scenario. The percent of changes correctly identified was calculated as the number of changes to which the participant responded with the correct classification, divided by the total number of changes.

Performance on the change detection task was collapsed across all experimental scenarios for periods of low and high demand separately. Repeated measure t-tests were conducted on both measures of change detection performance to determine the effect of the task demand manipulation. As expected, a significant effect of task demand was found for the percent of changes detected [$t(59) = 15.004$, $p < .001$], with performance significantly better under low demand ($M = 78.74\%$, $SD = 13.01$) than high demand ($M = 62.49\%$, $SD = 18.69$, $d = 1.01$). The same effect occurred for the percent of changes correctly identified [$t(59) = 18.328$, $p < .001$], with performance again significantly better under low demand ($M = 61.97\%$, $SD = 11.63$) than high demand ($M = 43.86\%$, $SD = 12.74$, $d = 1.48$).

## Performance

### Change Detection Performance

Both the percent of changes detected and the percent of changes correctly identified were evaluated through mixed-model ANOVAs using a 2 x 2 structure with variables *type of automation* (driving or beep alerts, between subjects) and *automation adaptability* (static or adaptive, within subjects). Detailed descriptive statistics are provided in Appendix F, Table 3.

#### Percent of Changes Detected

Significant main effects were found for type of automation [$F(1, 58) = 50.519$, $p < .001$] and automation adaptability [$F(1, 58) = 79.166$, $p < .001$]. The participants who received the beep alert automation performed significantly better ($M = 77.95\%$, $SD = 12.49$) than those receiving the driving automation ($M = 55.02\%$, $SD = 12.49$, $d = 1.84$). Participants also performed significantly better in scenarios with static automation ($M = 70.96\%$, $SD = 14.27$) than adaptive automation ($M = 62.00\%$, $SD = 11.80$, $d = 0.684$).

A significant interaction between automation adaptability and type of automation was also found [$F(1, 58) = 47.663$, $p < .001$]. This effect was further evaluated by examining the effect of automation adaptability within each type of automation separately (Figure 17). Within the beep alert automation condition, a significant main effect for automation adaptability was found [$F(1, 30) = 113.850$, $p < .001$], with the static automation scenarios ($M = 85.91\%$, $SD = 15.57$) performing better than the adaptive scenarios ($M = 69.99\%$, $SD = 12.79$, $d = 1.12$). Within the driving automation condition, the main effect for automation adaptability was not significant [$F(1, 28) = 2.247$, $p = .145$].

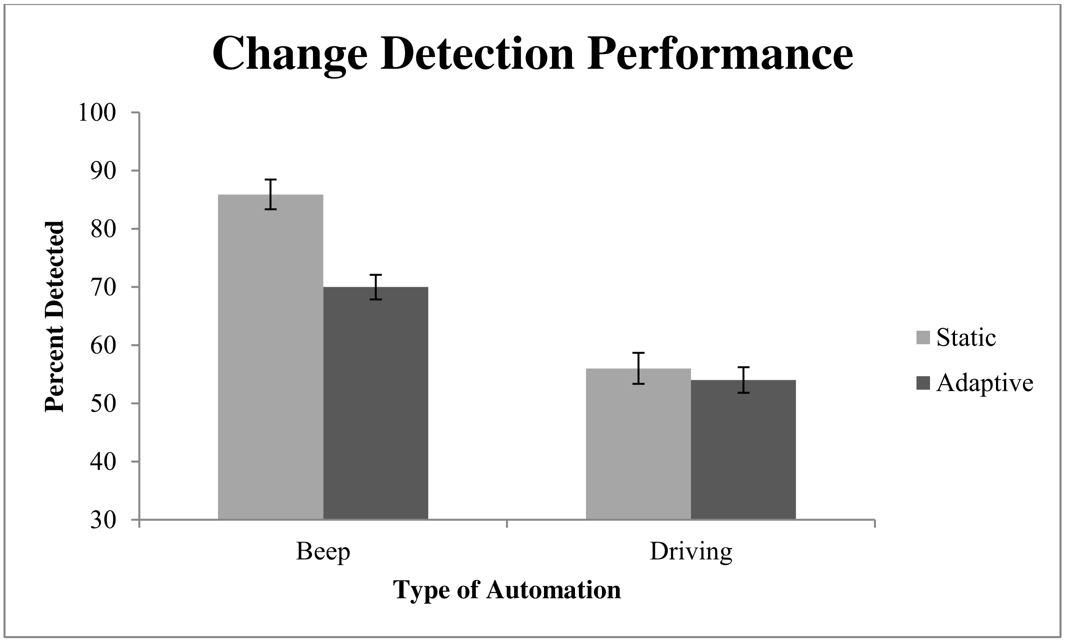


**Figure 17.** Percent of changes detected as a function of automation type and adaptability.

**Percent of Changes Correctly Identified**

Significant main effects were found for type of automation [$F(1, 58) = 24.720$, $p < .001$] and automation adaptability [$F(1, 58) = 56.398$, $p < .001$]. The participants who received the beep alert automation performed significantly better ($M = 55.02\%$, $SD = 10.12$) than those receiving the driving automation ($M = 42.03\%$, $SD = 10.12$, $d = 1.28$). Participants also performed significantly better in scenarios with static automation ($M = 51.52\%$, $SD = 11.12$) than adaptive automation ($M = 45.53\%$, $SD = 10.02$, $d = 0.566$).

A significant interaction between automation adaptability and type of automation was also found [$F(1, 58) = 18.551$, $p < .001$]. This effect was further evaluated by examining the effect of automation adaptability within each type of automation separately (Figure 18). Within the beep alert automation condition, a significant main effect for automation adaptability was found [$F(1, 30) = 85.003$, $p < .001$], with the static automation scenarios ($M = 59.73\%$, $SD = 11.25$) performing better than the adaptive scenarios ($M = 50.31\%$, $SD = 10.60$, $d = 0.862$). Within the driving automation condition, the main effect for automation adaptability was significant, though less pronounced [$F(1, 28) = 4.275$, $p = .048$]. Again, the static automation scenarios ($M = 43.31\%$, $SD = 10.96$) were found to perform better than the adaptive scenarios ($M = 40.75\%$, $SD = 9.354$, $d = 0.251$).

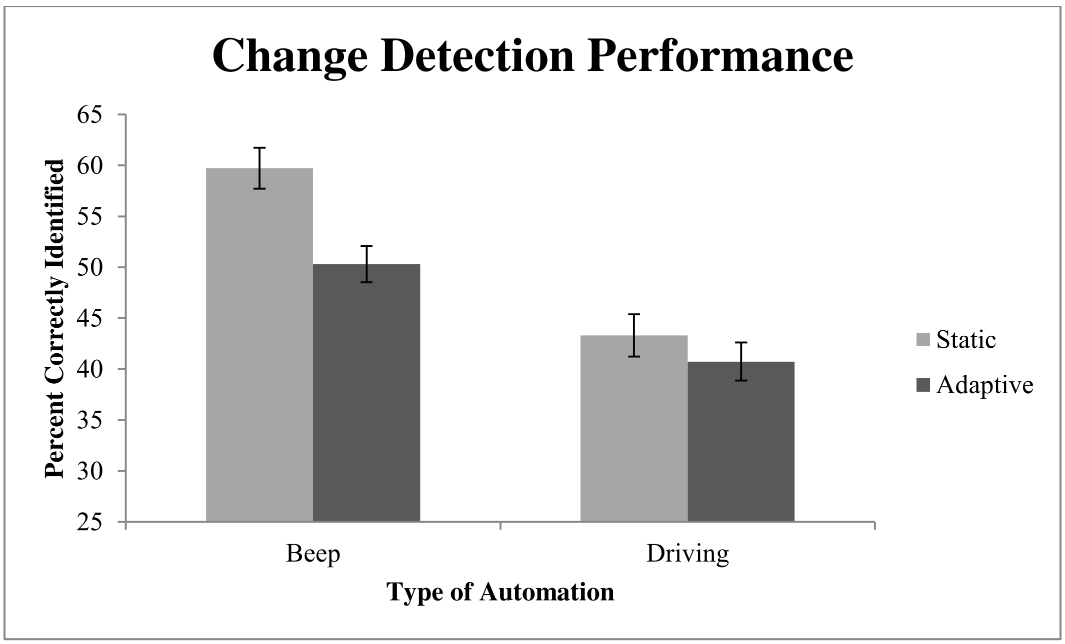


**Figure 18.** Percent of changes correctly identified as a function of automation type and adaptability.

**Personality Moderation**

Responses to the IPIP questionnaire were used to calculate values for Extraversion and Neuroticism for each participant. These values were tested for a potential moderating effect between automation adaptability and performance on the change detection task using General Linear Model analysis. Both Neuroticism and Extraversion failed to exhibit a significant main effect on either of the change detection performance metrics, and no significant moderating effects were found between automation adaptability and either of the personality traits ($p > .05$ in each case).

## Threat Detection Performance

Performance on the threat detection task was evaluated using metrics developed from signal detection theory. Signal detection theory traditionally evaluates performance in terms of sensitivity and bias. Sensitivity is the ability to discriminate signals (threats) from non-signals (friendlies), and bias is the tendency to be lenient, conservative, or unbiased when determining if a signal is present. Both measures are calculated based on hit rate (the percent of threats correctly detected) and false alarm rate (percent of friendlies incorrectly classified as threats).

A relatively high number of instances where no false alarms were reported required the use of nonparametric indices of sensitivity and bias. The index *A'* was calculated to measure sensitivity (Craig, 1979; Szalma, Hancock, Warm, Dember, & Parsons, 2006) and $\beta_D''$ for bias (See, Warm, Dember, & Howe, 1997; Szalma et al., 2006) using the formulas shown in Figure 19. As with change detection performance, both of these values were evaluated through mixed-model ANOVAs using a 2 x 2 structure with variables *type of automation* (driving or beep alerts, between subjects) and *automation adaptability* (static or adaptive, within subjects). Detailed descriptive statistics are provided in Appendix F, Table 4.

$$A' = 1 - .25\left[\frac{1-h}{1-f} + \frac{f}{h}\right]$$

$$\beta_D'' = \frac{(1-h)(1-f) - (h)(f)}{(1-h)(1-f) + (h)(f)}$$

**Figure 19.** Formulas used to calculate sensitivity (*A'*) and bias ($\beta_D''$), where *h* is hit rate and *f* is false alarm rate.

**Sensitivity**

Automation adaptability was found to have a significant effect on sensitivity [$F(1, 56) = 17.510$, $p < .001$], with sensitivity higher with adaptive automation ($M = 0.938$, $SD = 0.0286$) than static ($M = 0.922$, $SD = 0.0382$, $d = 0.474$). The main effect for type of automation and the interaction were not statistically significant ($p > .05$ in each case; Figure 20).

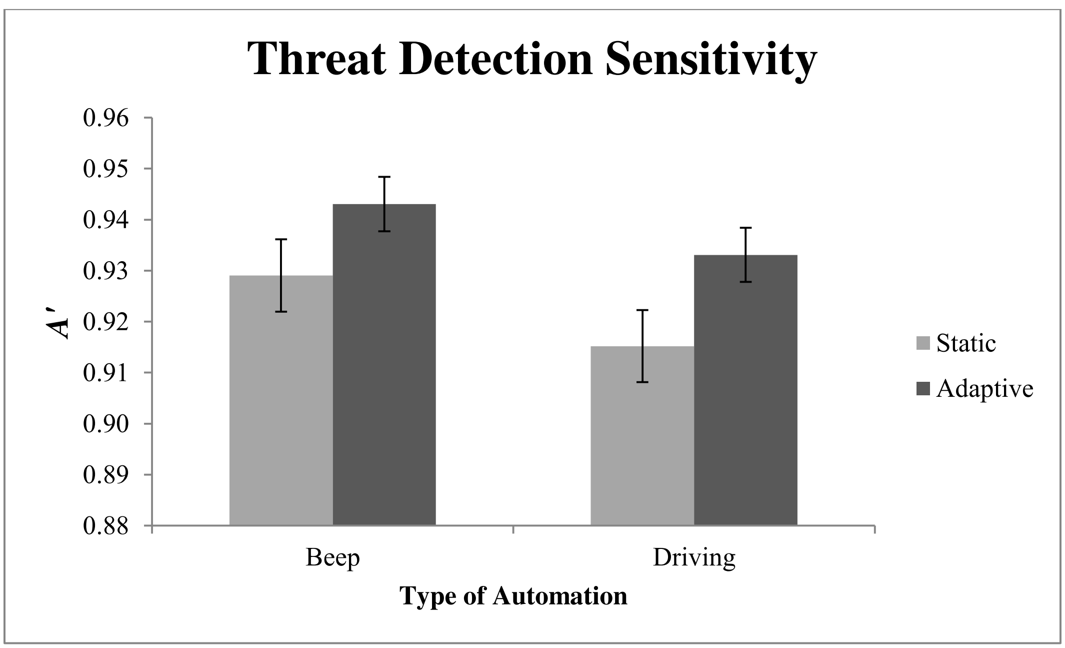


**Figure 20.** Sensitivity when detecting threats as a function of automation type and adaptability.

**Bias**

Type of automation was found to have a significant effect on bias [$F(1, 56) = 5.460$, $p = .023$], with bias higher (more conservative) with beep automation ($M = 1.000$, $SD = 0.0845$) than driving automation ($M = 0.948$, $SD = 0.0845$, $d = 0.615$).

Automation adaptability was found to have a significant effect on bias [$F(1, 56) = 5.328$, $p = .025$], with bias higher (more conservative) with adaptive automation ($M = 0.998$, $SD = 0.0067$) than static ($M = 0.950$, $SD = 0.163$, $d = 0.416$).

The type of automation x automation adaptability interaction was also found to be significant [$F(1, 56) = 5.328$, $p = .025$]. This effect was further evaluated by examining the effect of automation adaptability within each type of automation separately (Figure 21). Within the driving automation condition, a significant main effect for automation adaptability was found [$F(1, 28) = 5.328$, $p = .029$], with the adaptive automation scenarios ($M = 0.996$, $SD = 0.011$) higher (more conservative) than the static scenarios ($M = 0.900$, $SD = 0.232$, $d = 0.585$). Within the beep alert automation condition, the main effect for automation adaptability was not significant.

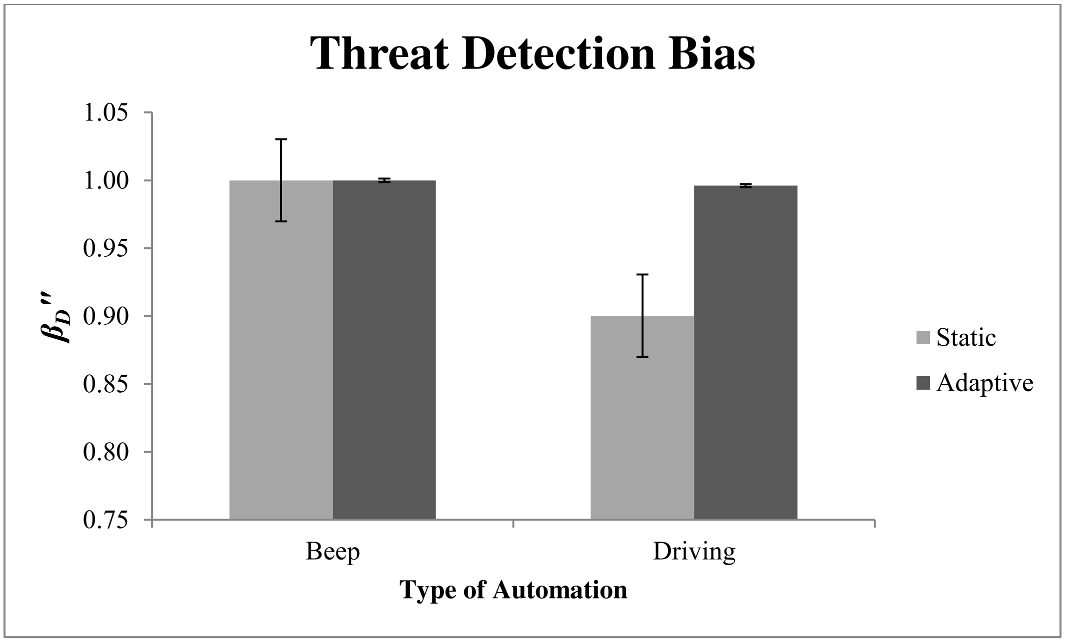

**Figure 21.** Response bias on the threat detection task.

### Personality Moderation

Responses to the IPIP questionnaire were used to calculate values for Extraversion and Neuroticism for each participant. These values were tested for a potential moderating effect between automation adaptability and performance on the threat detection task using General Linear Model analysis. Although no significant moderating effects were found between automation adaptability and either of the personality traits ($p > .05$ in each case), a significant negative relationship was found between Neuroticism and sensitivity [$F(1, 55) = 4.254$, $p = .044$] with those higher in Neuroticism having lower sensitivity on the threat detection task.

## Questionnaires

### Stress (DSSQ)

Responses to the DSSQ were used to calculate values of Distress, Engagement, and Worry (Figure 22). This data was first used to evaluate the global effects of the experimental task using repeated-measure *t*-tests to compare the pre-test measures to the average value computed across all four experimental scenarios. Detailed descriptive statistics are provided in Appendix F, Table 5. Results indicated that performance of the experimental task caused a significant increase in Distress [$M$ (difference) = 2.88, $SD$ = 5.21; $t(59) = -4.278$, $p < .001$] and decrease in Engagement [$M$ (difference) = 5.36, $SD$ = 4.75; $t(59) = 8.752$, $p < .001$] relative to pre-task values, with no significant effect on Worry ($p = .081$).

The values computed from the pre-task administration were then subtracted from the values obtained from each experimental scenario to account for individual variation in baseline stress. The resulting change scores were each evaluated through mixed-model ANOVAs using a 2 x 2 structure with variables *type of automation* (driving or beep alerts, between subjects) and *automation adaptability* (static or adaptive, within subjects).

The results indicated a significant main effect for type of automation on Worry [$F(1, 58) = 4.465$, $p = .039$]. The participants who received the beep alert automation reported significantly lower levels of Worry ($M = -2.685$, $SD = 5.26$) than those who received the driving automation ($M = 0.250$, $SD = 5.38$, $d = 0.552$). No other significant main effects or interactions were found ($p > .05$ in each case).

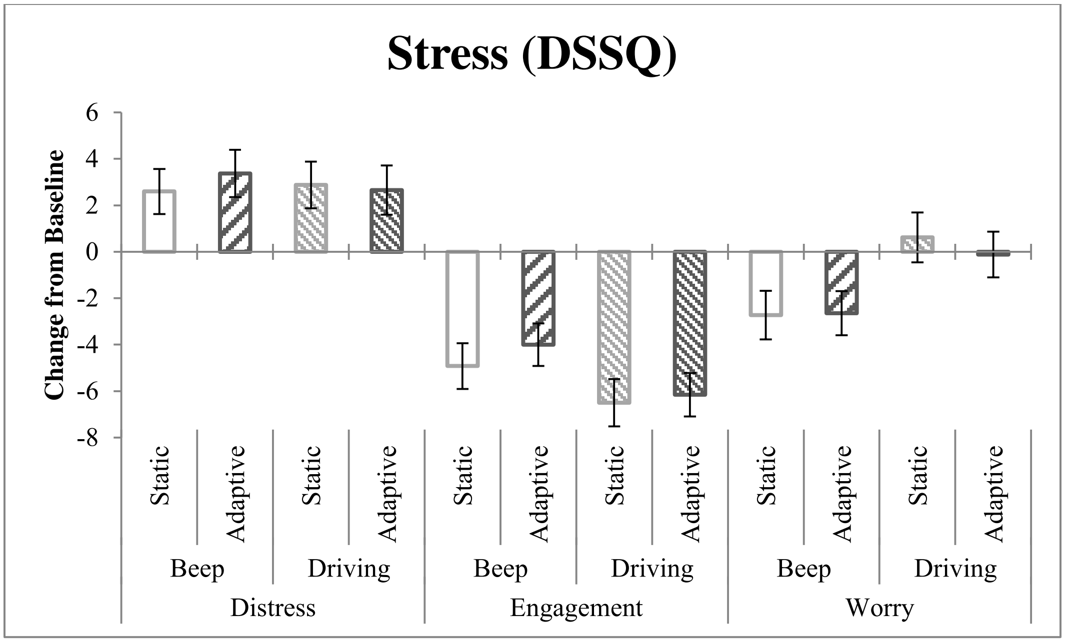


**Figure 22.** Stress reported from DSSQ responses as a function of type and adaptability of automation.

### Workload (NASA-TLX)

The NASA-TLX produced six workload subscales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration Level, as well as a single combined measure of global Workload based on the weighted average of the six subscales. Each of these values were evaluated through mixed-model ANOVAs using a 2 x 2 structure with variables *type of automation* (driving or beep alerts, between subjects) and *automation adaptability* (static or adaptive, within subjects). The sample used for these analyses is reduced due to missing data from one participant in the beep alert condition (Figure 23). Detailed descriptive statistics are provided in Appendix F, Table 6.

### Temporal Demand

A significant main effect of automation type was found for the Temporal Demand subscale [$F(1, 57) = 6.395$, $p = .014$]. The participants who received the beep alert automation reported significantly higher levels of Temporal Demand ($M = 64.92$, $SD = 21.34$) than those who received the driving automation ($M = 50.86$, $SD = 21.34$, $d = 0.659$).

### Effort

A significant main effect of automation type was found for the Effort subscale [$F(1, 57) = 10.235$, $p = .002$]. The participants who received the beep alert automation reported significantly higher levels of Effort ($M = 73.96$, $SD = 14.84$) than those who received the driving automation ($M = 61.60$, $SD = 14.84$, $d = 0.833$).

### Performance

A significant main effect was found for automation adaptability on the Performance scale [$F(1, 57) = 6.721$, $p = .012$]. Participants rated this scale higher (indicating that they believed their performance was worse) for scenarios with static automation ($M = 60.95$, $SD = 21.08$) than those with adaptive automation ($M = 56.08$, $SD = 21.62$, $d = 0.228$).

### Frustration, Mental Demand, and Physical Demand

No significant main effects or interactions were found for the Frustration, Mental Demand, or Physical Demand subscales ($p > .05$ in each case).

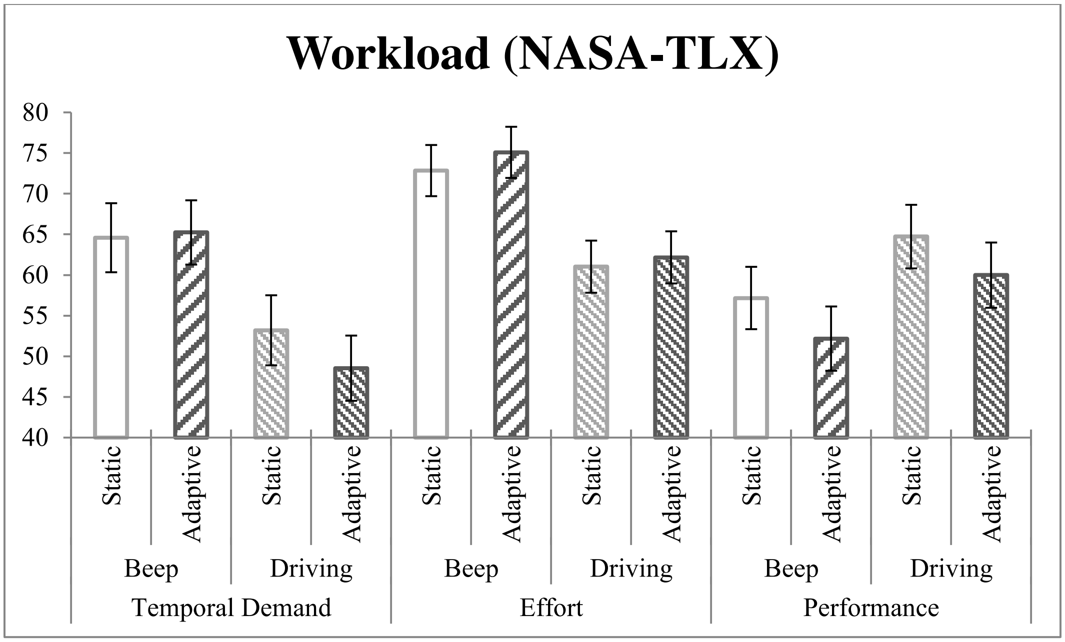

**Figure 23.** Workload reported from NASA-TLX responses as a function of type and adaptability of automation.

## Physiological Measures

### Electroencephalogram (EEG)

Data collected from the EEG was calculated using the engagement index (Freeman et al., 1999; Pope et al., 1995). Rather than evaluating separate bands of EEG activity independently, activity across three common bands is combined to form a single value [beta / (alpha + theta)]. Data from sensor sites Cz, Pz, P3, and P4 was used, with each individual's baseline value subtracted from their activity during the scenario to produce a change from baseline value. This data was evaluated through a mixed-model ANOVA using a 2 x 2 structure with variables *type of automation* (driving or beep alerts, between subjects) and *automation adaptability* (static or adaptive, within subjects). The sample used for these analyses is reduced to 55 (28 beep automation, 27 driving automation) due to errors in physiological data collection. Results from this analysis yielded no significant main effects or interactions ($p > .05$ in each case). Detailed descriptive statistics are provided in Appendix F, Table 7.

**Electrocardiogram (ECG)**

Data collected from the ECG was used to determine heart rate variability (HRV), which is the statistical variance of the time period between heartbeats. Heartbeats are initially detected using the So and Chan method (Tan, Chan, & Choi, 2000). Baseline HRV is subtracted from the values calculated for each experimental scenario to account for individual variation. This data was evaluated through a mixed-model ANOVA using a 2 x 2 structure with variables *type of automation* (driving or beep alerts, between subjects) and *automation adaptability* (static or adaptive, within subjects). The sample used for these analyses is reduced to 57 (30 beep automation, 27 driving automation) due to errors in physiological data collection. Detailed descriptive statistics are provided in Appendix F, Table 8.

Type of automation was found to have a significant effect on HRV [$F(1, 55) = 5.336$, $p = .025$]. Those receiving the beep automation had lower HRV values ($M = 0.610$, $SD = 22.41$) than those receiving the driving automation ($M = 14.34$, $SD = 22.41$, $d = 0.613$).

The interaction between type of automation and adaptability was also significant [$F(1, 55) = 11.518$, $p = .001$]. Further analysis revealed an effect of adaptability on HRV, the direction of which changed with automation type (Figure 24). For those participants who received the beep alert automation [$F(1, 29) = 6.159$, $p = .019$], adaptive automation resulted in higher HRV values ($M = 3.287$, $SD = 21.72$) than static automation ($M = -2.068$, $SD = 22.31$, $d = 0.243$). This trend was reversed for those who received the driving automation [$F(1, 26) = 5.396$, $p = .028$], with static automation resulting in higher HRV values (*M* = 17.519, *SD* = 28.30) than adaptive automation (*M* = 11.160, *SD* = 20.51, *d* = 0.257).

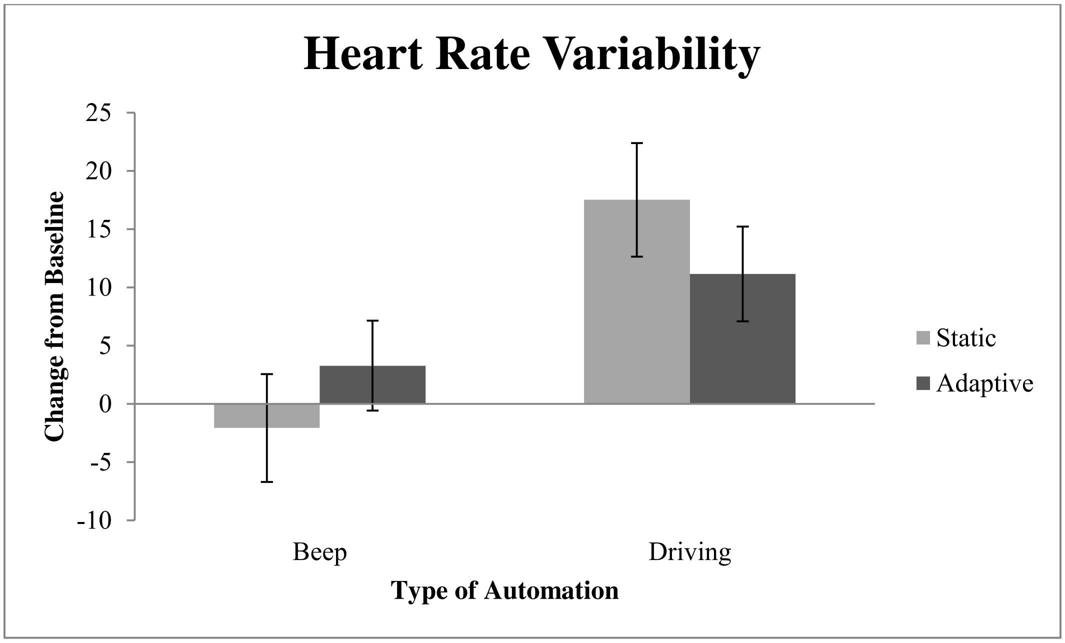


**Figure 24.** HRV as a function of adaptability and type of automation. Lower values indicate higher levels of workload.

# DISCUSSION

## Hypothesis H1

### Summary of Results

The proposed hypothesis was partially supported by the results. Hypothesis H1, which predicted that demand-type matched automation would improve performance, workload (decreasing Mental and Temporal Demand), and stress (decreasing Distress), was supported by the beep alert automation condition performing significantly better than the driving automation condition on both measures of change detection performance. Although no effect of automation type on the Worry dimension of stress was predicted, those in the beep alert automation condition also reported significantly lower levels of Worry than those in the driving automation condition. However, no significant difference was found between the two automation conditions on their performance of the threat detection task, or the Distress and Task Engagement dimensions of stress.

Automation type was found to have a significant effect on ratings of Temporal Demand, but in a direction opposite from that predicted, with those in the beep alert automation condition reporting higher levels of Temporal Demand than those in the driving automation condition. Similarly, automation type also had a significant effect on HRV, but in an unexpected direction, with those in the beep alert automation condition having lower HRV values (indicating higher levels of workload) than those in the driving automation condition, indicating a performance-workload dissociation caused by the automation type manipulation. Although no effect of automation type on the Effort dimension of workload was predicted, those in the beep alert automation condition also reported significantly higher levels of Effort than those in the driving automation condition.

**Discussion**

As predicted, the beep alert automation, designed to specifically address the specific demands the operator is subjected to by the task, did improve their performance. Although this improvement in performance was limited to the change detection task, the lack of a decrease in performance on the threat detection task indicates that a net improvement in overall operator performance occurred, rather than a shifting of focus from one task to another. However, it was predicted that performance would improve as a result of the freeing of cognitive resources, which did not occur. In fact, the beep alert automation was found to significantly increase the Temporal Demand and Effort scales of the NASA-TLX, and decrease HRV (indicating an increase in workload), resulting in a performance-workload dissociation (Yeh & Wickens, 1988).

This joint effect of improving performance while also increasing workload could be clarified through two potential explanations. First, the task demand could be such that the participants performing them fall on the lower end of the curvilinear relationship between workload and performance, the hypostress region of dynamic instability in Hancock and Warm's (1989) model. Therefore, an increase in workload would be expected to elicit a corresponding increase in performance. However, given the magnitude of the values reported on the various NASA-TLX subscales, this seems unlikely. Although entirely subjective, further evidence against this theory is the fact that participants frequently mentioned that they found the task to be particularly difficult, indicating that they were experiencing higher than average levels of workload.

Yeh and Wickens (1988) suggest that performance-workload dissociations are often the result of the investment of greater resources to the performance of a resource-limited task. Therefore, the results are more effectively explained through the cognitive-energetical model (Hockey, 1997; Hockey, Gaillard, & Coles, 1986). This model does not aim to reject alternative resource-based theories, but rather proposes the addition of compensatory effort. Hockey suggests that an operator's performance on a task is not simply based on the level of workload they experience, but also on the actions of a higher level, goal-focused managerial system. This system maintains goals for both performance and cognitive/emotional well-being (i.e. workload and stress), and is capable of making deliberate sacrifices in one area to benefit the other (Figure 25). For example, when an operator is met with an increase in task difficulty, they can respond by allowing their performance to decline to maintain their cognitive state, or increase their effort (subjecting themselves to greater workload and stress) to maintain performance.

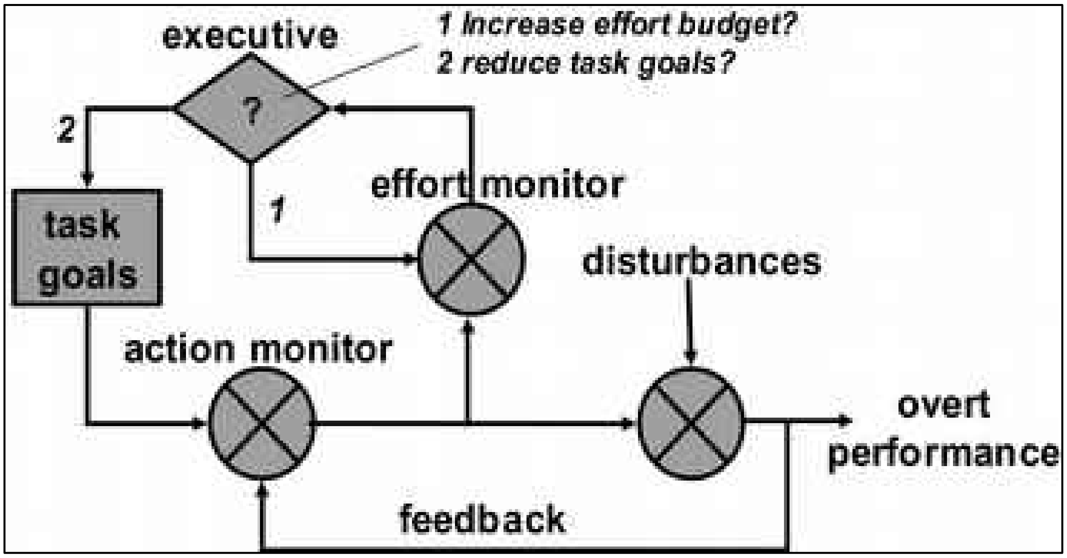


**Figure 25.** Hockey's cognitive-energetical model of compensatory effort (Hockey, 1997).

Within this model, levels of both performance and cognitive energy devoted to the task are monitored continuously through feedback loops. These levels are each compared to their corresponding goal states, and the decision to adjust the level of effort devoted to the task is based on the discrepancy between the current value and goal state, with each discrepancy weighted based on the relative importance of its associated goal state. It is believed to be through this mechanism that the beep automation causes its simultaneous increase in performance and workload. This is achieved through the increase in saliency of change detection events caused by the beep alert. By making each signal more salient, the operator can more easily recognize any time they miss responding to one. Increasing their awareness of missed signals will inherently cause a corresponding decrease in their perceived performance. Therefore, given two operators with equivalent performance, one receiving the beep alert automation will likely believe they performed worse than one receiving the driving automation. By decreasing the operator's self-perceived level of performance, the discrepancy from their performance goal state is increased. This increased discrepancy motivates them to sacrifice additional cognitive energy to elevate their perceived performance level closer to their goal state, explaining the beep alert automation condition's higher levels of change detection performance, workload (measured by HRV), and subjective Effort and Temporal Demand. This theory also explains why performance was significantly better for those who received the beep alert automation, but subjective performance ratings (from the NASA-TLX subscale) were equivalent across automation conditions.

The one effect of the automation manipulation not directly explained by the cognitive-energetical model is the fact that those who received the beep alert automation reported significantly lower levels of the Worry dimension of stress than those who received the driving automation. Given that Worry is representative of the cognitive processes of stress (including self-focus, self-esteem, task-related cognitive interference and task-irrelevant cognitive interference), it is usually found to decline with time on task. This decline is typically most prevalent in the self-focus and task-irrelevant cognitive interference facets, as focus shifts away from the self and is devoted to the task (Matthews et al., 1999). This pattern is consistent across many types of tasks, including reading, card sorting, and working memory tasks (Matthews et al., 2002), and is evident from the participants who received the beep alert automation. However, the level of Worry reported from those receiving the driving automation remained unchanged from baseline values, a trend typically only found from the performance of visual vigilance tasks. Therefore, it appears that the implementation of the driving automation changes the structure of the task in a way that causes it to become a sustained attention task (although other hallmarks of vigilance, such as reduced sensitivity and engagement over time, were not evident). This may result from the fact that the driving task is the only continuous-control portion of the experimental scenarios. Therefore, removing this task from the operator through automation leaves them with only the threat detection and change detection tasks to perform, both of which are fundamentally signal detection processes. Sheridan (1992) specifically discusses this issue as a potential pitfall of the use of automation in the realm of robotic control tasks, and these findings offer further support for his claims.

## Hypothesis H2

### Summary of Results

Hypothesis H2, which predicted that adapting the level of automation to the level of task demand would improve performance and stress (increasing Task Engagement and decreasing Distress), was supported by the adaptive automation scenarios performing better than the static automation scenarios on the threat detection task. Hypothesis H2 was also supported by the ECG data, which found the scenarios with static automation to have lower HRV scores than the adaptive automation scenarios, indicating that adaptive automation reduced the level of workload (for the beep alert condition only). Additionally, although no effect of automation adaptability was predicted for the Performance workload subscale, ratings for adaptive scenarios were significantly lower (indicating that participants felt they performed better) than static automation scenarios.

However, support for this hypothesis had several limitations. Threat detection bias was only reduced by static automation for those who received the driving automation. Further, the effect of static automation having lower HRV scores was limited to the beep alert automation condition, with participants in the driving automation condition demonstrating the opposite pattern.

The expected effect of automation adaptability on the stress dimensions were not present, with no significant effect found for either Task Engagement or Distress. The effect of automation adaptability on change detection performance was reversed from expectations, with static automation resulting in better performance than adaptive, with this effect stronger for those in the beep alert automation condition than the driving automation condition. The effect on HRV was also reversed from expectations for the driving automation condition, with the adaptive scenarios resulting in lower HRV values (indicating higher levels of workload) than the static scenarios.

**Discussion**

The hypothesis was partially supported through the HRV findings, with scenarios in which adaptive beep alert automation was provided resulting in lower levels of workload than the static beep alert automation. It was predicted that a similar trend would follow for the driving automation, though to a lesser extent, but in fact those who received the driving automation actually exhibited higher levels of workload under adaptive automation conditions than static. This provides considerable support for the primary research question: that matching the type of automation to the type of demand being experienced is critical, particularly within an adaptive environment.

However, the use of adaptive automation does not come without a cost. Adaptive automation led to poorer change detection performance than static automation regardless of the type of automation, though the effect was stronger with the beep alert automation than the driving automation. For the beep alert automation, this effect is likely attributable to the nature of this automation. The fact that the beep alert automation elicits greater effort from the operator to improve their performance suggests that it is likely impossible to reduce workload without also reducing performance when using this form of automation. However, the implementation of adaptive driving automation caused a decrease in performance while also causing an increase in workload (relative to static driving automation). This is likely caused by the driving automation doing little to alleviate the mental demands imposed by the task. Although teleoperation of a remote vehicle can be mentally demanding under certain conditions, the vehicle operation task for this experiment was intentionally designed to elicit a low level of demand by using simple routes, with only eight turns distributed across 24 minutes and no obstacles or interference. Therefore, adaptively manipulating the level of the driving automation provides no benefit, with the fluctuations in the task environment serving only as a distraction (Reinerman-Jones, Taylor, et al., 2011).

## Hypothesis H3

### Summary of Results

Hypothesis H3, which predicted that individual differences in personality (Extraversion and Neuroticism) would moderate the influence of automation adaptability on performance and stress, received no empirical support. Although no significant moderating effects were found for either of the personality dimensions, a significant relationship was found between Neuroticism and threat detection, with those higher in Neuroticism having lower sensitivity.

## Discussion

Although the personality measures failed to moderate the effects of automation type or adaptability, the effect of Neuroticism on threat detection sensitivity indicates that the task is sensitive to individual differences. The lack of moderation effects may simply be due to insufficient statistical power, as individual differences effects are traditionally fairly weak, requiring substantial sample sizes to find their effects (Szalma, 2009). Szalma and Taylor (2011) also found evidence that providing highly reliable automation may attenuate the detrimental impact Neuroticism has on performance, and so the use of perfectly reliable automation in the current study may have further reduced the strength of any potential moderating effects. Although unrelated to the automation manipulations, the results did provide additional support for prior evidence (Szalma & Taylor, 2011) that those higher in Neuroticism perform worse on threat detection tasks.

# Conclusions

The use of the driving automation, unmatched to the type of demand subjected by the task, provided relatively little benefit to the operator, and in fact showed evidence of disengaging them from the task. However, the use of the beep alert automation, designed to specifically support the cognitive faculties under the greatest demand, significantly improved mission performance. Further, when the level of automation varied adaptively based on the level of demand imposed by the task, operator workload was reduced when provided with the beep alert automation, but workload actually increased for operators using the driving automation.

However, the use of the beep alert automation was still not ideal. The beep alert was intended to offload some of the perceptual demands of the task by increasing the saliency of the change events. This proved to be an effective method, resulting in improved change detection performance. However, the beep alert also caused a simultaneous increase in workload (Effort and Temporal Demand, specifically). The beep alert automation appeared to improve operator performance not through the alleviation of cognitive demands, as was expected, but rather by motivating the operator to sacrifice additional cognitive energy by making them more aware of their performance errors. Although ultimately successful in its primary goal of improving performance, this associated cost in operator cognitive resources is an important factor to consider before such an aid is implemented in any system.

The theoretical implications of this study demonstrate that the type of automation implemented within an environment has a considerable impact on the operator, in terms of performance as well as their cognitive/emotional state. These results contradict previous theory which proposed that humans are best supported by automation of the action implementation phase of information processing (Kaber et al., 2005). It appears that such a generalized statement is not true across all task types, but rather the type of automation which best supports the operator is that which supports the cognitive dimension most burdened by their task. Providing automation which does not support the appropriate cognitive dimension can result in the many potential problems with automation (supervisory control, disengagement, skill degradation, etc.) without accomplishing any of its intended benefits.

**Future Research**

Although the current study provides preliminary support for the importance of matching automation type to the type of demand experienced by the operator, additional research is necessary to ensure that this is true for all types of demand. The task used in the current study focused only on subjecting (and alleviating) perceptual demands. Therefore, further evaluation is necessary to evaluate the same concept under high levels of other types of demand, such as decision making or action implementation.

Also, before a complex system can become truly adaptive to various types of demand, more diagnostic measures of specific types of cognitive demand must be developed. Most metrics of cognitive state derived from physiological measures that could feasibly be implemented in a real-world setting still classify workload along a single continuum, incapable of discriminating between various types of mental demand. However, metrics intended to classify general cognitive states, such as the Engagement Index, may be determined to in fact be measuring specific subcomponents of cognitive activity. In this study, the Engagement Index failed to detect differences between task manipulations of demand or automation type, despite these manipulations having significant effects on performance, as well as other subjective and physiological measures of operator state. This suggests that the Engagement Index may be sensitive only to specific cognitive functions, which does not necessarily invalidate its potential utility, but demonstrates the necessity for further research to determine exactly what cognitive functions it is capable of measuring. These measures must evolve dramatically before a system can be capable of truly understanding the operator's cognitive state on a multidimensional level in real-time, a necessary capability before the system can adapt to meet the operator's specific needs.

Finally, additional research is also needed to better understand the influence of individual differences (e.g. personality traits) on performance and cognitive/emotional state within an adaptive system. Prior research has suggested that personality plays an important role in how an operator interacts with automated systems, but the current study failed to find any evidence of this. If true, future systems can utilize knowledge of the operator's traits, in addition to their fluctuating states, to better meet their specific needs. However, larger sample sizes will be necessary to have sufficient statistical power before a definitive conclusion can be made regarding the effect of these individual differences.

## Application

This study provides further support for the multidimensionality of cognitive resources, and demonstrates the importance of considering these dimensions when implementing automation. Although this study evaluated the importance of demand-type matched automation within a military UGV control setting, the findings are in no way limited to the operation of unmanned vehicles, or even military tasks. Any complex task environment, in which more than one type of demand may be experienced by the operator, would benefit from matching automation type to the demand type currently experienced by the operator. In fact, designers of even relatively simple tasks, in which only a single form of demand is present, must also consider whether the automated assistance they provide to the operator truly supports the demand imposed by the task. This study provides further evidence that the traditional "automate what you can" model of system design fails to support the operator. Serious consideration must be given to the implementation of automation if any benefit is to come of it; failure to do so risks employing automation which provides little to no operational advantage, or worse, actually impairs the operator's ability to perform their task.

# APPENDIX A: DEMOGRAPHICS QUESTIONNAIRE

## Demographics Questionnaire

Participant # _______ Age ______ Major ________________ Date ___________ Gender ___

1. What is the <u>highest</u> level of education you have had?

Less than 4 yrs of college ____ Completed 4 yrs of college ____ Other ____

2. When did you use computers in your education? (*<u>Circle all that apply</u>*)

| | | |
|---|---|---|
| Grade School | Jr. High | High School |
| Technical School | College | Did Not Use |

3. Where do you currently use a computer? (*<u>Circle all that apply</u>*)

Home Work Library Other________ Do Not Use

4. For each of the following questions, <u>circle</u> the response that best describes you.

| How often do you: | |
|---|---|
| Use a mouse? | Daily, Weekly, Monthly, Once every few months, Rarely, Never |
| Use a joystick? | Daily, Weekly, Monthly, Once every few months, Rarely, Never |
| Use a touch screen? | Daily, Weekly, Monthly, Once every few months, Rarely, Never |
| Use icon-based programs/software? | Daily, Weekly, Monthly, Once every few months, Rarely, Never |
| Use programs/software with pull-down menus? | Daily, Weekly, Monthly, Once every few months, Rarely, Never |
| Use graphics/drawing features in software packages? | Daily, Weekly, Monthly, Once every few months, Rarely, Never |
| Use E-mail? | Daily, Weekly, Monthly, Once every few months, Rarely, Never |
| Operate a radio controlled vehicle (car, boat, or plane)? | Daily, Weekly, Monthly, Once every few months, Rarely, Never |
| Play computer/video games? | Daily, Weekly, Monthly, Once every few months, Rarely, Never |

5. Which type(s) of computer/video games do you most often play if you play at least once every few months?

6. Which of the following best describes your expertise with computers? (check √ one)
   _____ Novice
   _____ Good with one type of software package (such as word processing or slides)
   _____ Good with several software packages
   _____ Can program in one language and use several software packages
   _____ Can program in several languages and use several software packages

7. Are you in your usual state of health physically? YES NO
   If NO, please briefly explain:

8. How many hours of sleep did you get last night? ______ hours

9. Do you have normal color vision? YES NO

10. Do you have prior military service? YES NO If Yes, how long __________

11. Are you currently serving in the military? YES NO
    If yes, are you off duty at the time you are participating in this study? YES NO

# APPENDIX B: PERSONALITY MEASURE

**IPIP Questionnaire**

On the following pages, there are phrases describing people's behaviors. Please use the rating scale below to describe how accurately each statement describes ***you***. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age.

So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence. Please read each statement carefully, and then circle the number that corresponds to the value on the scale:

| Very Inaccurate | Moderately Inaccurate | Neither Inaccurate nor Accurate | Moderately Accurate | Very Accurate |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

| | | | | | |
|---|---|---|---|---|---|
| 1. I would describe my experiences as somewhat dull. | 1 | 2 | 3 | 4 | 5 |
| 2. I don't talk a lot. | 1 | 2 | 3 | 4 | 5 |
| 3. I have frequent mood swings. | 1 | 2 | 3 | 4 | 5 |
| 4. I make friends easily. | 1 | 2 | 3 | 4 | 5 |
| 5. I don't like to draw attention to myself. | 1 | 2 | 3 | 4 | 5 |
| 6. I have little to say. | 1 | 2 | 3 | 4 | 5 |
| 7. I feel comfortable around people. | 1 | 2 | 3 | 4 | 5 |
| 8. I keep in the background. | 1 | 2 | 3 | 4 | 5 |
| 9. I seldom feel blue. | 1 | 2 | 3 | 4 | 5 |

| | | | | | |
|---|---|---|---|---|---|
| 10. I dislike myself. | 1 | 2 | 3 | 4 | 5 |
| 11. I often feel blue. | 1 | 2 | 3 | 4 | 5 |
| 12. I am skilled in handling social situations. | 1 | 2 | 3 | 4 | 5 |
| 13. I rarely get irritated. | 1 | 2 | 3 | 4 | 5 |
| 14. I feel comfortable with myself. | 1 | 2 | 3 | 4 | 5 |
| 15. I am not easily bothered by things. | 1 | 2 | 3 | 4 | 5 |
| 16. I panic easily. | 1 | 2 | 3 | 4 | 5 |
| 17. I am the life of the party. | 1 | 2 | 3 | 4 | 5 |
| 18. I know how to captivate people. | 1 | 2 | 3 | 4 | 5 |
| 19. I am often down in the dumps. | 1 | 2 | 3 | 4 | 5 |
| 20. I am very pleased with myself. | 1 | 2 | 3 | 4 | 5 |

# APPENDIX C: DSSQ PRE-TEST

## DSSQ: Pre-test

<u>General Instructions</u>

This questionnaire is concerned with your feelings and thoughts at the moment. Please answer **every** question, even if you find it difficult. Answer, as honestly as you can, what is true of you. Please do not choose a reply just because it seems like the 'right thing to say'. Your answers will be kept entirely confidential. Also, be sure to answer according to how you feel **AT THE MOMENT**. Don't just put down how you usually feel. You should try and work quite quickly: there is no need to think very hard about the answers. The first answer you think of is usually the best.

For each statement, circle an answer from 0 to 4, so as to indicate how accurately it describes your feelings **AT THE MOMENT**.

**Definitely false = 0, Somewhat false = 1,**
**Neither true nor false = 2, Somewhat true = 3, Definitely true = 4**

| | | | | | |
|---|---|---|---|---|---|
| 1. The content of the task will be dull. | 0 | 1 | 2 | 3 | 4 |
| 2. I feel relaxed | 0 | 1 | 2 | 3 | 4 |
| 3. I am determined to succeed on the task. | 0 | 1 | 2 | 3 | 4 |
| 4. I feel tense. | 0 | 1 | 2 | 3 | 4 |
| 5. Generally, I feel in control of things. | 0 | 1 | 2 | 3 | 4 |
| 6. I am reflecting about myself. | 0 | 1 | 2 | 3 | 4 |
| 7. My attention is directed towards the task. | 0 | 1 | 2 | 3 | 4 |
| 8. I am thinking deeply about myself. | 0 | 1 | 2 | 3 | 4 |
| 9. I feel energetic. | 0 | 1 | 2 | 3 | 4 |
| 10. I am thinking about something that happened earlier today. | 0 | 1 | 2 | 3 | 4 |
| 11. I will find the task too difficult for me. | 0 | 1 | 2 | 3 | 4 |
| 12. I will find it hard to keep my concentration on the task. | 0 | 1 | 2 | 3 | 4 |
| 13. I am thinking about personal concerns and interests. | 0 | 1 | 2 | 3 | 4 |
| 14. I feel confident about my performance. | 0 | 1 | 2 | 3 | 4 |
| 15. I am examining my motives. | 0 | 1 | 2 | 3 | 4 |
| 16. I feel like I could handle any difficulties I encounter. | 0 | 1 | 2 | 3 | 4 |
| 17. I am motivated to try hard at the task. | 0 | 1 | 2 | 3 | 4 |
| 18. I am thinking about things important to me. | 0 | 1 | 2 | 3 | 4 |
| 19. I feel uneasy. | 0 | 1 | 2 | 3 | 4 |
| 20. I feel tired. | 0 | 1 | 2 | 3 | 4 |

# APPENDIX D: DSSQ POST-TEST

## DSSQ: Post-test

General Instructions

This questionnaire is concerned with your feelings and thoughts while you were performing the task. Please answer **every** question, even if you find it difficult. Answer, as honestly as you can, what is true of you. Please do not choose a reply just because it seems like the 'right thing to say'. Your answers will be kept entirely confidential. Also, be sure to answer according to how you felt **WHILE PERFORMING THE TASK**. Don't just put down how you usually feel. You should try and work quite quickly: there is no need to think very hard about the answers. The first answer you think of is usually the best.

For each statement, circle an answer from 0 to 4, so as to indicate how accurately it describes your feelings **WHILE PERFORMING THE TASK**.

**Definitely false = 0, Somewhat false = 1,**
**Neither true nor false = 2, Somewhat true = 3, Definitely true = 4**

| | | | | | |
|---|---|---|---|---|---|
| 1. The content of the task was dull. | 0 | 1 | 2 | 3 | 4 |
| 2. I felt relaxed. | 0 | 1 | 2 | 3 | 4 |
| 3. I was determined to succeed on the task. | 0 | 1 | 2 | 3 | 4 |
| 4. I felt tense. | 0 | 1 | 2 | 3 | 4 |
| 5. Generally, I felt in control of things. | 0 | 1 | 2 | 3 | 4 |
| 6. I reflected about myself. | 0 | 1 | 2 | 3 | 4 |
| 7. My attention was directed towards the task. | 0 | 1 | 2 | 3 | 4 |
| 8. I thought deeply about myself. | 0 | 1 | 2 | 3 | 4 |
| 9. I felt energetic. | 0 | 1 | 2 | 3 | 4 |
| 10. I thought about something that happened earlier today. | 0 | 1 | 2 | 3 | 4 |
| 11. I found the task too difficult for me. | 0 | 1 | 2 | 3 | 4 |
| 12. I found it hard to keep my concentration on the task. | 0 | 1 | 2 | 3 | 4 |
| 13. I thought about personal concerns and interests. | 0 | 1 | 2 | 3 | 4 |
| 14. I felt confident about my performance. | 0 | 1 | 2 | 3 | 4 |
| 15. I examined my motives. | 0 | 1 | 2 | 3 | 4 |
| 16. I felt like I could handle any difficulties I encountered. | 0 | 1 | 2 | 3 | 4 |
| 17. I was motivated to try hard at the task. | 0 | 1 | 2 | 3 | 4 |
| 18. I thought about things important to me. | 0 | 1 | 2 | 3 | 4 |
| 19. I felt uneasy. | 0 | 1 | 2 | 3 | 4 |
| 20. I felt tired. | 0 | 1 | 2 | 3 | 4 |

# APPENDIX E: NASA-TLX

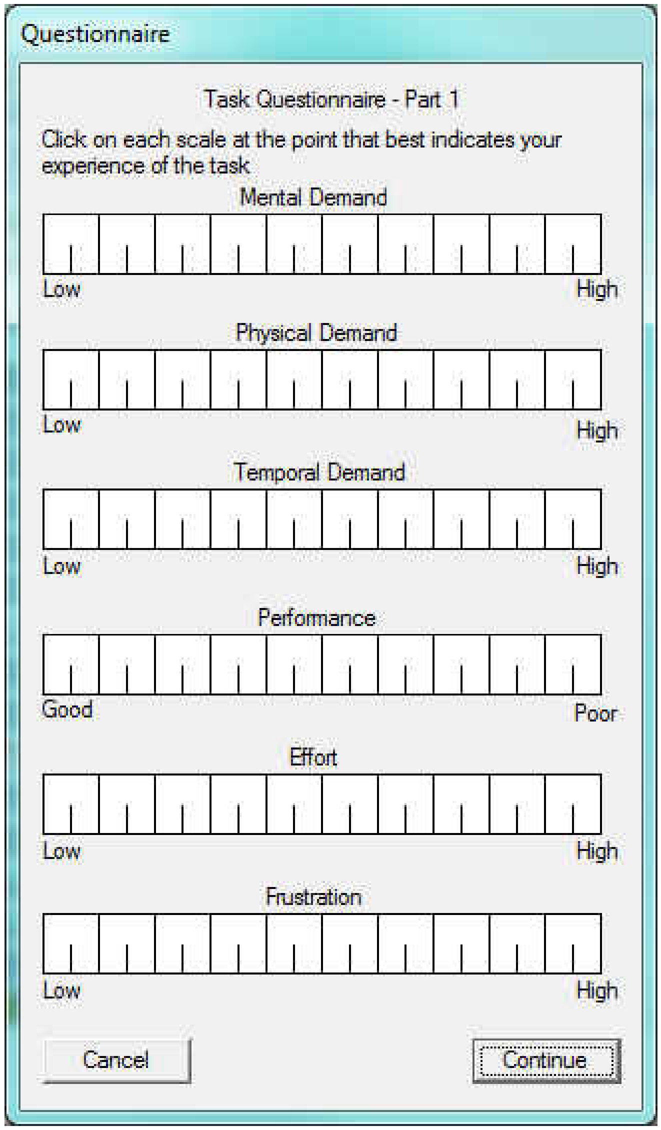


**Figure 26.** Part 1 of the NASA-TLX computer program. The participant uses a mouse to indicate their rating of each scale.

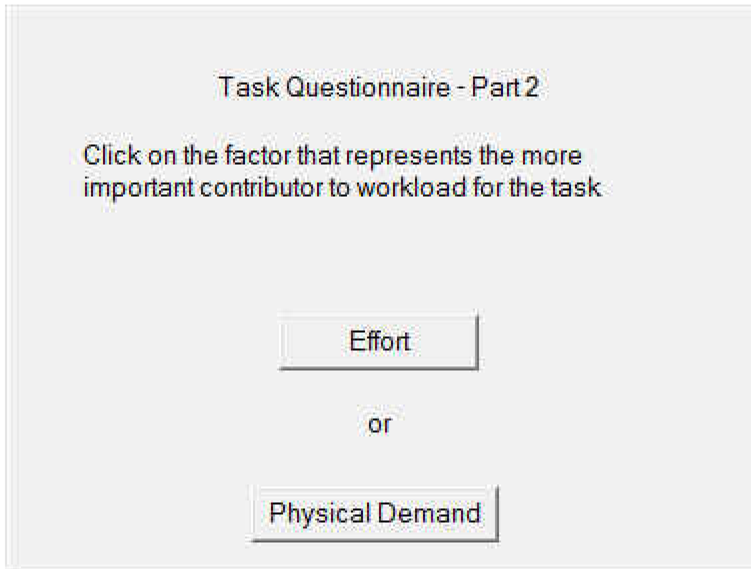


**Figure 27.** Part 2 of the NASA-TLX computer program. The participant is presented with all possible pair-wise comparisons of the six scales (a total of 15).

# APPENDIX F: DESCRIPTIVE STATISTICS

**Table 3.** Change detection values.

| Measure | Condition | Adaptability | N | Mean | Std. Error | Extraversion Correlation | | | Neuroticism Correlation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | *r* | *p* | N | *r* | *p* | N |
| Percent of Changes Detected | Beep | Static | 31 | 85.907 | 2.561 | -.075 | .693 | 30 | .130 | .493 | 30 |
| | | Adaptive | 31 | 69.991 | 2.118 | -.101 | .596 | 30 | .056 | .770 | 30 |
| | Driving | Static | 29 | 56.019 | 2.648 | -.132 | .495 | 29 | .070 | .720 | 29 |
| | | Adaptive | 29 | 54.011 | 2.190 | -.089 | .648 | 29 | -.185 | .337 | 29 |
| Percent of Changes Correctly Identified | Beep | Static | 31 | 59.732 | 1.995 | -.197 | .296 | 30 | .039 | .838 | 30 |
| | | Adaptive | 31 | 50.312 | 1.799 | -.143 | .450 | 30 | .025 | .896 | 30 |
| | Driving | Static | 29 | 43.305 | 2.063 | -.099 | .611 | 29 | .012 | .951 | 29 |
| | | Adaptive | 29 | 40.752 | 1.860 | -.136 | .481 | 29 | -.209 | .276 | 29 |

**Table 4.** Threat detection values.

| Measure | Condition | Adaptability | N | Mean | Std. Error | Extraversion Correlation | | | Neuroticism Correlation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | *r* | *p* | N | *r* | *p* | N |
| Sensitivity (*A'*) | Beep | Static | 29 | .929 | .007 | -.076 | .695 | 29 | -.406 | .029 | 29 |
| | | Adaptive | 29 | .943 | .005 | .238 | .214 | 29 | -.377 | .044 | 29 |
| | Driving | Static | 29 | .915 | .007 | .103 | .595 | 29 | -.145 | .452 | 29 |
| | | Adaptive | 29 | .933 | .005 | .118 | .543 | 29 | -.039 | .842 | 29 |
| Bias ($\beta_D''$) | Beep | Static | 29 | 1.000 | .000 | * | * | * | * | * | * |
| | | Adaptive | 29 | 1.000 | .000 | * | * | * | * | * | * |
| | Driving | Static | 29 | .900 | .030 | -.234 | .223 | 29 | -.085 | .660 | 29 |
| | | Adaptive | 29 | .996 | .001 | -.090 | .642 | 29 | -.096 | .620 | 29 |

* Correlations cannot compute because bias values are constant

**Table 5.** Stress (DSSQ) values. All values are reported as change from baseline.

| Measure | Condition | Adaptability | N | Mean | Std. Error | Extraversion Correlation | | | Neuroticism Correlation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | *r* | *p* | N | *r* | *p* | N |
| Distress | Beep | Static | 31 | 2.597 | .970 | -.012 | .951 | 30 | .206 | .274 | 30 |
| | | Adaptive | 31 | 3.371 | 1.020 | .000 | .999 | 30 | .082 | .667 | 30 |
| | Driving | Static | 29 | 2.879 | 1.003 | -.082 | .671 | 29 | -.136 | .481 | 29 |
| | | Adaptive | 29 | 2.655 | 1.055 | .140 | .470 | 29 | .073 | .708 | 29 |
| Engagement | Beep | Static | 31 | -4.919 | .988 | -.094 | .621 | 30 | .099 | .603 | 30 |
| | | Adaptive | 31 | -4.000 | .912 | -.073 | .702 | 30 | .051 | .789 | 30 |
| | Driving | Static | 29 | -6.500 | 1.021 | .170 | .377 | 29 | .195 | .310 | 29 |
| | | Adaptive | 29 | -6.155 | .942 | .017 | .931 | 29 | .155 | .423 | 29 |
| Worry | Beep | Static | 31 | -2.726 | 1.043 | -.199 | .292 | 30 | .327 | .078 | 30 |
| | | Adaptive | 31 | -2.645 | .952 | -.256 | .172 | 30 | .315 | .090 | 30 |
| | Driving | Static | 29 | .621 | 1.078 | -.446 | .015 | 29 | .256 | .180 | 29 |
| | | Adaptive | 29 | -.121 | .984 | -.287 | .131 | 29 | .199 | .301 | 29 |

**Table 6.** Workload (NASA-TLX) values.

| Measure | Condition | Adaptability | N | Mean | Std. Error | Extraversion Correlation | | | Neuroticism Correlation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | *r* | *p* | N | *r* | *p* | N |
| Total Workload | Beep | Static | 30 | 69.372 | 1.920 | .401 | .028 | 30 | -.059 | .757 | 30 |
| | | Adaptive | 30 | 69.328 | 1.896 | .361 | .050 | 30 | -.108 | .569 | 30 |
| | Driving | Static | 29 | 66.328 | 1.953 | .172 | .373 | 29 | .161 | .403 | 29 |
| | | Adaptive | 29 | 63.868 | 1.928 | .031 | .875 | 29 | .164 | .395 | 29 |
| Physical Demand | Beep | Static | 30 | 39.667 | 4.045 | .086 | .653 | 30 | .263 | .160 | 30 |
| | | Adaptive | 30 | 40.500 | 4.313 | .004 | .983 | 30 | .378 | .039 | 30 |
| | Driving | Static | 29 | 29.483 | 4.114 | .078 | .689 | 29 | -.383 | .040 | 29 |
| | | Adaptive | 29 | 32.328 | 4.387 | .042 | .829 | 29 | .006 | .976 | 29 |
| Temporal Demand | Beep | Static | 30 | 64.583 | 4.240 | .374 | .042 | 30 | .150 | .428 | 30 |
| | | Adaptive | 30 | 65.250 | 3.953 | .368 | .045 | 30 | .055 | .773 | 30 |
| | Driving | Static | 29 | 53.190 | 4.312 | .230 | .229 | 29 | .047 | .810 | 29 |
| | | Adaptive | 29 | 48.534 | 4.021 | .346 | .066 | 29 | .067 | .729 | 29 |
| Performance | Beep | Static | 30 | 57.167 | 3.848 | .345 | .062 | 30 | .041 | .829 | 30 |
| | | Adaptive | 30 | 52.167 | 3.948 | .211 | .263 | 30 | .121 | .525 | 30 |
| | Driving | Static | 29 | 64.741 | 3.914 | -.063 | .747 | 29 | .118 | .541 | 29 |
| | | Adaptive | 29 | 60.000 | 4.015 | -.309 | .103 | 29 | .122 | .529 | 29 |
| Effort | Beep | Static | 30 | 72.833 | 3.143 | .494 | .006 | 30 | -.211 | .262 | 30 |
| | | Adaptive | 30 | 75.083 | 3.151 | .184 | .332 | 30 | -.294 | .115 | 30 |
| | Driving | Static | 29 | 61.034 | 3.197 | .107 | .580 | 29 | .018 | .925 | 29 |
| | | Adaptive | 29 | 62.155 | 3.205 | .080 | .680 | 29 | -.066 | .732 | 29 |
| Frustration | Beep | Static | 30 | 53.833 | 4.545 | .333 | .072 | 30 | .193 | .308 | 30 |
| | | Adaptive | 30 | 53.917 | 4.364 | .173 | .361 | 30 | .260 | .165 | 30 |
| | Driving | Static | 29 | 55.345 | 4.622 | .185 | .336 | 29 | .002 | .992 | 29 |
| | | Adaptive | 29 | 52.069 | 4.438 | .318 | .093 | 29 | .039 | .841 | 29 |
| Mental Demand | Beep | Static | 30 | 81.917 | 2.488 | .005 | .977 | 30 | -.129 | .496 | 30 |
| | | Adaptive | 30 | 81.667 | 2.718 | .095 | .618 | 30 | -.109 | .567 | 30 |
| | Driving | Static | 29 | 80.172 | 2.530 | .030 | .877 | 29 | .258 | .177 | 29 |
| | | Adaptive | 29 | 78.362 | 2.764 | .075 | .698 | 29 | .216 | .260 | 29 |

**Table 7.** EEG Engagement Index values. All values are reported as change from baseline.

| Measure | Condition | Adaptability | N | Mean | Std. Error | Extraversion Correlation | | | Neuroticism Correlation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | *r* | *p* | N | *r* | *p* | N |
| EEG Engagement Index | Beep | Static | 28 | 1.549 | .385 | .211 | .264 | 30 | -.183 | .332 | 30 |
| | | Adaptive | 28 | 1.184 | 1.705 | .186 | .324 | 30 | -.309 | .097 | 30 |
| | Driving | Static | 27 | .866 | .392 | -.213 | .267 | 29 | .034 | .862 | 29 |
| | | Adaptive | 27 | 3.046 | 1.737 | .160 | .408 | 29 | -.051 | .795 | 29 |

**Table 8.** ECG Heart Rate Variability (HRV) values. All values are reported as change from baseline.

| Measure | Condition | Adaptability | N | Mean | Std. Error | Extraversion Correlation | | | Neuroticism Correlation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | *r* | *p* | N | *r* | *p* | N |
| ECG HRV | Beep | Static | 30 | -2.068 | 4.623 | .013 | .946 | 29 | .077 | .690 | 29 |
| | | Adaptive | 30 | 3.287 | 3.862 | -.129 | .504 | 29 | .158 | .413 | 29 |
| | Driving | Static | 27 | 17.519 | 4.873 | .178 | .375 | 27 | .008 | .970 | 27 |
| | | Adaptive | 27 | 11.160 | 4.071 | .217 | .276 | 27 | .064 | .752 | 27 |

# REFERENCES


Abarbanel, A. (1999). The neural underpinnings of neurofeedback training. In J. R. Evans & A. Abarbanel (Eds.), *Introduction to quantitative EEG and neurofeedback* (pp. 311-340). San Diego, CA: Academic Press.

Amalberti, R. (1999). Automation in aviation: A human factors perspective. In D. J. Garland, J. A. Wise, & D. Hopkin (Eds.), *Handbook of Aviation Human Factors* (pp. 173-192). Mahwah, NJ: Lawrence Erlbaum Associates.

Bailey, N. R., Scerbo, M. W., Freeman, F. G., Mikulka, P. J., & Scott, L. A. (2006). Comparison of a brain-based adaptive system and a manual adaptable system for invoking automation. *Human Factors*, *48*(4), 693-709. doi:10.1518/001872006779166280

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775-779. doi:10.1016/0005-1098(83)90046-8

Barber, D., Leontyev, S., Sun, B., Davis, L., Nicholson, D., & Chen, J. Y. C. (2008). The Mixed Initiative Experimental (MIX) testbed for collaborative human robot interactions. *Army Science Conference*. Orlando, FL.

Barnes, M., Parasuraman, R., & Cosenzo, K. A. (2006). *Adaptive automation for military robotic systems. In NATO Technical Report RTOTR- HFM-078 Uninhabited military vehicles: Human factors issues in augmenting the force* (pp. 420-440).

Bliss, J. (1997). Alarm reaction patterns by pilots as a function of reaction modality. *The International Journal of Aviation Psychology*, *7*(1), 1-14. Taylor & Francis. doi:10.1207/s15327108ijap0701_1

Byrne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology*, *42*(3), 249-268.

Campbell-Kelly, M., & Aspray, W. (2004). *Computer: A History of the Information Machine*. New York, NY: Westview Press.

Chambers, A. B., & Nagel, D. C. (1985). Pilots of the future: Human or computer? *Communications of the ACM*, *28*(11), 1187-1199.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual*. *Odessa FL Psychological Assessment Resources*. Psychological Assessment Resources.

Cox, T. (1978). *Stress*. London, England: Macmillan.

Cox-Fuenzalida, L. E., Swickert, R., & Hittner, J. B. (2004). Effects of neuroticism and workload history on performance. *Personality and Individual Differences*, *36*(2), 447-456. Elsevier. doi:10.1016/S0191-8869(03)00108-9

Craig, A. (1979). Nonparametric measures of sensory efficiency for sustained monitoring tasks. *Human Factors*, *21*(1), 69-77.

Cummings, M. L., & Guerlain, S. (2007). Developing operator capacity estimates for supervisory control of autonomous vehicles. *Human Factors*, *49*(1), 1-15.

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. New York, NY: Academic Press.

Dearden, A., Harrison, M., & Wright, P. (2000). Allocation of function: Scenarios, context and the economics of effort. *International Journal of Human-Computer Studies*, *52*(2), 289-318. doi:10.1006/ijhc.1999.0290

Department of Defense. (2005). *Department of Defense interface standard: Common Warfighter symbology (MIL-STD-2525B)*.

Dornheim, M. A. (1999). Apache tests power of new cockpit tool. *Aviation Week & Space Technology*, *151*(16), 46-69.

Edwards, E. (1977). Automation in civil transport aircraft. *Applied Ergonomics*, *8*(4), 194-198. doi:10.1016/0003-6870(77)90163-6

Eggemeier, F. T., Wilson, G. F., Kramer, A. F., & Damos, D. (1991). Workload assessment in multi-task environments. In D. L. Damos (Ed.), *Multiple-task performance* (pp. 207-216). London, England: Taylor & Francis.

Entin, E. E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors*, *41*(2), 312-325. doi:10.1518/001872099779591196

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. New York, NY: Plenum Press.

Fitts, P. M. (1951). *Human engineering for an effective air-navigation and traffic-control system*. Columbus, OH: National Research Council, Division of Anthropology and Psychology, Committee on Aviation Psychology.

Freeman, F. G., Mikulka, P. J., Pope, A. T., Prinzel, L. J., & Scerbo, M. W. (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. SAGE Publications.

Freeman, F. G., Mikulka, P. J., Prinzel, L. J., & Scerbo, M. W. (1999). Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology*, *50*(1), 61-76. doi:10.1016/S0301-0511(99)00002-2

Funke, G., Matthews, G., Warm, J. S., & Emo, A. K. (2007). Vehicle automation: A remedy for driver stress? *Ergonomics*, *50*(8), 1302-1323. doi:10.1080/00140130701318830

Goldberg, L., Johnson, J., Eber, H., Hogan, R., Ashton, M., Cloninger, C., & Gough, H. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*, 84-96. doi:10.1016/j.jrp.2005.08.007

Gopher, D., & Donchin, E. (1986). Workload – An examination of the concept. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance (Vol 2: Cognitive processes and performance)* (pp. 41-1-41-49). New York, NY: John Wiley & Sons, Inc.

Groover, M. P. (2007). *Automation, Production Systems, and Computer-integrated Manufacturing* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Hancock, P. A., & Chignell, M. H. (1987). Adaptive control in human-machine systems. In P. A. Hancock (Ed.), *Human factors psychology* (pp. 305-345). Amsterdam: North-Holland.

Hancock, P. A., & Parasuraman, R. (1992). Human factors and safety in the design of intelligent vehicle-highway systems (IVHS). *Journal of Safety Research*, *23*(4), 181-198. doi:10.1016/0022-4375(92)90001-P

Hancock, P. A., & Scallen, S. F. (1996). The Future of Function Allocation. *Ergonomics in Design: The Quarterly of Human Factors Applications*, *4*(4), 24-29.

Hancock, P. A., & Szalma, J. L. (2003). Operator stress and display design. *Ergonomics in Design: The Quarterly of Human Factors Applications*, *11*(2), 13-18. SAGE Publications.

Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors*, *31*(5), 519-537.

Hancock, P. A., Oron-Gilad, T., & Szalma, J. L. (2007). Elaborations of the multiple resource theory of attention. In A. F. Kramer, D. A. Wiegmann, & A. Kirlik (Eds.), *Attention: From theory to practice* (pp. 45-56). Oxford, England: Oxford University Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=wEUi-bEaqWcC&oi=fnd&pg=PA45&dq=Elaborations+of+the+multiple+resource+theory+of+attention&ots=51DkW2HC3p&sig=1_p9V7g9dcT70F4Q1cNiR2he2s4

Hancock, P. A., Parasuraman, R., & Byrne, E. A. (1996). Driver-centered issues in advanced automation for motor vehicle. In R. Parasuraman & M. Mouloua (Eds.), *Automation and*

*Human Performance: Theory and Applications* (pp. 337-364). Mahwah, NJ: Lawrence Erlbaum Associates.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-184). North-Holland: Elsevier Science Publishers.

Hebb, D. O. (1955). Drives and the C.N.S. (conceptual nervous system). *Psychological Review*, *62*, 243-254.

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. *Proceedings of the Seventh International Symposium on Aviation Psychology1* (pp. 161-166). Columbus, OH: The Department of Aviation, The Avitation Psychology Laboratory, The Ohio State University.

Hitchcock, E., Warm, J. S., Matthews, G., Dember, W. N., Shear, P., Tripp, L., Mayleben, D., et al. (2003). Automation cueing modulates cerebral blood flow and vigilance in a simulated air traffic control task. *Theoretical Issues in Ergonomics Science*, *4*(1), 89-112. doi:10.1080/14639220210159726

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, *45*(1-3), 73-93. doi:10.1016/S0301-0511(96)05223-4

Hockey, G. R. J., Gaillard, A. W. K., & Coles, M. G. H. (Eds.). (1986). *Energetics and human information processing*. Dordrecht, The Netherlands: Martinus Nijhoff.

Hollnagel, E., & Woods, D. D. (1999). Cognitive Systems Engineering: New wine in new bottles. *International Journal of Human-Computer Studies*, *51*(2), 339-356. doi:10.1006/ijhc.1982.0313

Hopkin, V. D. (1992). Human factors issues in air traffic control. *Human Factors Society Bulletin*, *35*(6), 1-4.

Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science*, *5*(2), 113-153.

Kaber, D. B., & Riley, V. (1999). Adaptive automation of a dynamic control task based on workload assessment through a secondary monitoring task. In M W Scerbo & M. Mouloua (Eds.), *Automation Technology and Human Performance: Current Research and Trends* (pp. 129-133). Mahwah, NJ: Erlbaum.

Kaber, D. B., Omal, E., & Endsley, M. R. (1999). Level of automation effects on telerobot performance and human operator situation awareness and subjective workload. In M W Scerbo (Ed.), *Automation Technology and Human Performance: Current Research and Trends* (pp. 165-170). Mahwah, N: Erlbaum.

Kaber, D. B., Perry, C., Segall, N., McClernon, C., & Prinzel, L. J. (2006). Situation awareness implications of adaptive automation for information processing in an air traffic control-related task. *International Journal of Industrial Ergonomics*, *36*(5), 447-462. doi:10.1016/j.ergon.2006.01.008

Kaber, D. B., Wright, M. C., & Sheik-Nainar, M. A. (2006). Investigation of multi-modal interface features for adaptive automation of a human–robot system. *International Journal of Human-Computer Studies*, *64*(6), 527-540. doi:10.1016/j.ijhcs.2005.11.003

Kaber, D. B., Wright, M. C., Prinzel, L. J., & Clamann, M. P. (2005). Adaptive automation of human-machine system information-processing functions. *Human Factors*, *47*(4), 730-741. doi:10.1518/001872005775570989

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice Hall.

Kantowitz, B. H., & Knight, J. L. (1976). Testing tapping timesharing, II: Auditory secondary task. *Acta Psychologica*, *40*(5), 343-362. doi:10.1016/0001-6918(76)90016-0

Kramer, A. F., & Weber, T. (2000). Application of psychophysiology to human factors. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 794-814). New York, NY: Cambridge University Press.

Kramer, A. F., Sirevaag, E. J., & Braune, R. (1987). A psychophysiological assessment of operator workload during simulated flight missions. *Human Factors*, *29*(2), 145-160.

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford, England: Oxford University Press.

Lee, J., & Moray, N. E. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies*, *40*(1), 153-184. doi:10.1006/ijhc.1994.1007

Lee, J., & See, K. (2004). Trust in automation: designing for appropriate reliance. *Human Factors*, *46*(1), 50-80.

Liu, D., Wasson, R., & Vincenzi, D. A. (2008). Effects of system automation management strategies and multi-mission operator-to-vehicle ratio on operator performance in UAV systems. *Journal of Intelligent and Robotic Systems*, *54*(5), 795-810. doi:10.1007/s10846-008-9288-4

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *The Quarterly Journal of Experimental Psychology*, *1*(1), 6-21. doi:10.1080/17470214808416738

Matthews, G., Campbell, S., Falconer, S., Joyner, L., Huggins, J., Gilliland, K., Grier, R., et al. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion*, *2*(4), 315-340.

Matthews, G., Deary, I. J., & Whiteman, M. C. (2003). *Personality traits* (2nd ed.). Cambridge, MA: Cambridge University Press.

Matthews, G., Jones, D. M., & Chamberlain, A. G. (1992). Predictors of individual differences in mail-coding skills and their variation with ability level. *Journal of Applied Psychology*, *77*(4), 406-418. doi:10.1037/0021-9010.77.4.406

Matthews, G., Joyner, L., Gilliland, K., Campbell, S., Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire: Towards a state "Big Three"? In I. Mervielde, I. J. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe, Vol. 7* (Vol. 7, pp. 335-350). Tilburg, Netherlands: Tilburg University Press.

Merritt, S. M., & Ilgen, D. R. (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors*, *50*(2), 194-210.

Miller, C. A., & Hannen, M. D. (1999). The Rotorcraft Pilot's Associate: Design and evaluation of an intelligent user interface for cockpit information management. *Knowledge-Based Systems*, *12*(8), 443-456.

Mitchell, P. J., Cummings, M. L., & Sheridan, T. B. (2004). *Human supervisory control issues in network centric warfare (HAL2004-01)*. Cambridge, MA: Massachusetts Institute of Technology.

Moray, N. E. (1979). *Mental workload: Its theory and measurement*. New York, NY: Plenum Press.

Mosier, K. L., Skitka, L. J., & Korte, K. J. (1994). Cognitive and social psychological issues in flight crew/automation interaction. In M. Mouloua & R. Parasuraman (Eds.), *Human Performance in Automated Systems: Current Research and Trends* (pp. 191-197). Hillsdale, NJ: Erlbaum.

Mouloua, M., Gilson, R., & Koonce, J. (1997). Automation, flight management and pilot training: Issues and considerations. In R. A. Telfer & P. J. Moore (Eds.), *Aviation Training: Learners, Instruction and Organization* (pp. 78-86). Aldershot, United Kingdom: Avebury Aviation.

National Transportation Safety Board. (1973). *Aircraft Accident Report: Eastern Air Lines, Inc. L-1011, N310EA, Miami, Florida, 20 December 1972 (Rep. NTSB-AAR-73-14).* Washington, D.C.

Netherlands Aviation Safety Board. (1992). *Aircraft Accident Report 92-11 El Al Flight 1862 Boeing 747-258F 4X-AXG*. Bijlmermeer, Amsterdam.

Noyes, J. M. (2009). Vigilance and human supervisory control – A potted history. In P. D. Bust (Ed.), *Contemporary Ergonomics* (pp. 218-225). Wiltshire, UK: Taylor & Francis.

Opperman, R. (1994). *Adaptive user support*. Hillsdale, NJ: Erlbaum.

Parasuraman, R. (1986). Vigilance, monitoring, and search. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance, Vol. 2: Cognitive processes and performance* (pp. 1-39). Oxford, England: John Wiley and Sons.

Parasuraman, R. (1987). Human-computer monitoring. *Human Factors*, *29*(6), 695-706. Human Factors and Ergonomics Society.

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

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230-253. Human Factors and Ergonomics Society. doi:10.1518/001872097778543886

Parasuraman, R., Bahri, T., Molloy, R., & Singh, I. (1991). Effects of shifts in the level of automation on operator performance. *Proceedings of the Sixth International Symposium on Aviation Psychology* (pp. 102-107). Columbus, OH: The Department of Aviation, The Avitation Psychology Laboratory, The Ohio State University.

Parasuraman, R., Cosenzo, K. A., & De Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. *Military Psychology*, *21*(2), 270-297. doi:10.1080/08995600902768800

Parasuraman, R., Hancock, P. A., & Olofinboba, O. (1997). Alarm effectiveness in driver-centred collision-warning systems. *Ergonomics*, *40*(3), 390-399.

Parasuraman, R., Mouloua, M., & Hilburn, B. (1999). Adaptive aiding and adaptive task allocation enhance human-machine interaction. In M W Scerbo & M. Mouloua (Eds.), *Automation Technology and Human Performance: Current Research and Trends* (pp. 119-123). Mahwah, NJ: Erlbaum.

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

Parasuraman, R., Mouloua, M., Molloy, R., & Hilburn, B. (1993). Adaptive function allocation reduces performance costs of static automation. *Proceedings of the Seventh International Symposium on Aviation Psychology*. Columbus, OH: The Department of Aviation, The Avitation Psychology Laboratory, The Ohio State University.

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 - Part A: Systems and Humans*, *30*(3), 286-297. doi:10.1109/3468.844354

Parasuraman, R., Warm, J. S., & Dember, W. N. (1987). Vigilance: Taxonomy and utility. In L. S. Mark, J. S. Warm, & R. L. Huston (Eds.), *Ergonomics and human factors: Recent research* (pp. 11-32). New York, NY: Springer-Verlag.

Pitts, J. (2009, July). Ground robotics: Driving the road of the future. *Army AL&T*. Retrieved from http://www.usaasc.info/alt_online/article.cfm?iID=0907&aid=12

Poor, H. V. (1994). *An introduction to signal detection and estimation* (2nd ed.). New York, NY: Springer.

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology*, *40*(1-2), 187-195. doi:10.1016/0301-0511(95)05116-3

Prinzel, L. J., Parasuraman, R., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2003). *Three experiments examining the use of electroencephalogram, event-related potentials, and heart-rate variability for real- time human-centered adaptive automation design (NASA/TP-2003-212442)*. Hanover, MD.

Reinerman-Jones, L. E., Barber, D., Lackey, S., & Nicholson, D. (2010). Developing methods for utilizing physiological measures. *Applied Human Factors and Ergonomics Society Conference*. Miami, FL.

Reinerman-Jones, L. E., Matthews, G., Langheim, L., & Warm, J. S. (2011). Selection for vigilance assignments: A review and proposed new direction. *Theoretical Issues in Ergonomics Science*, *12*(4), 273-296. Taylor & Francis. doi:10.1080/14639221003622620

Reinerman-Jones, L. E., Taylor, G. S., Sprouse, K., Barber, D., & Hudson, I. (2011). Adaptive automation as a task switching and task congruence challenge. *Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society*. Las Vegas, NV.

Riley, V. (1994). A theory of operator reliance on automation. In M. Mouloua & R. Parasuraman (Eds.), *Human Performance in Automated Systems: Current Research and Trends* (pp. 8-14). Hillsdale, NJ: Erlbaum.

Riley, V. (1996). What avionics engineers should know about pilots and automation. *IEEE Aerospace and Electronic Systems Magazine*, *11*(5), 3-8. doi:10.1109/62.494182

Rouse, W. B. (1977). Human-computer interaction in multitask situations. *IEEE Transactions on Systems, Man, and Cybernetics*, *7*(5), 384-392. doi:10.1109/TSMC.1977.4309727

Rouse, W. B. (1988). Adaptive aiding for human/computer control. *Human Factors*, *30*, 431-443.

Rouse, W. B., Cannon-Bowers, J. A., & Salas, E. (1992). The role of mental models in team performance in complex systems. *IEEE Transactions on Systems, Man, and Cybernetics*, *22*(6), 1296-1308. doi:10.1109/21.199457

Saqer, H., De Visser, E., Emfield, A., Shaw, T., & Parasuraman, R. (2011). Adaptive automation to improve human performance in supervision of multiple uninhabited aerial vehicles: Individual markers of performance. *Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society* (pp. 890-893). Las Vegas, NV. doi:10.1177/1071181311551185

Scerbo, Mark W. (2001). Adaptive automation. In W. Karwoski (Ed.), *International encyclopedia of ergonomics and human factors* (pp. 1077-1079). London, England: Taylor and Francis, Inc.

See, J. E., Howe, S. R., Warm, J. S., & Dember, W. N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin*, *117*(2), 230-249. doi:10.1037/0033-2909.117.2.230

See, J. E., Warm, J. S., Dember, W. N., & Howe, S. R. (1997). Vigilance and signal detection theory: An empirical evaluation of five measures of response bias. *Human Factors*, *39*(1), 14-29.

Sheridan, T. B. (1992). *Telerobotics, Automation, and Human Supervisory Control*. Cambridge, MA: MIT Press.

Sheridan, T. B. (1997). Supervisory control. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (2nd ed., pp. 1295-1327). Hoboken, NJ: John Wiley and Sons.

Sheridan, T. B. (2000). Function allocation: Algorithm, alchemy or apostasy? *International Journal of Human-Computer Studies*, *52*(2), 203-216. doi:10.1006/ijhc.1999.0285

Sheridan, T. B. (2002). *Humans and automation: System design and research issues*. New York, NY: John Wiley.

Sheridan, T. B., & Verplank, W. L. (1978). *Human and computer control of undersea teleoperators*. Cambridge, MA: Cambridge University Press.

Smith, C., & Lazarus, R. S. (1990). Emotion and adaptation. In L. A. Pervin (Ed.), *Handbook of Personality: Theory and Research* (pp. 609-637). New York, NY: Guilford.

Squire, P., Trafton, G., & Parasuraman, R. (2006). Human control of multiple unmanned vehicles: Effects of interface type on execution and task switching times. *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction* (pp. 26-32).

Stanton, N. A., & Young, M. S. (1998). Vehicle automation and driving performance. *Ergonomics*, *41*(7), 1014-1028. doi:10.1080/001401398186568

Szalma, J. L. (2008). Individual differences in stress reaction. In P. A. Hancock & J. L. Szalma (Eds.), *Performance under stress* (pp. 323-358). Hampshire, England: Ashgate Publishing.

Szalma, J. L. (2009). Individual differences in human–technology interaction: Incorporating variation in human characteristics into human factors and ergonomics research and design. *Theoretical Issues in Ergonomics Science*, *10*(5), 381-397. doi:10.1080/14639220902893613

Szalma, J. L., & Hancock, P. A. (2007). Task loading and stress in human-computer interaction: Theoretical frameworks and mitigation strategies. In A. Sears & J. Jacko (Eds.), *Handbook for Human-Computer Interaction in Interactive Systems* (2nd ed., pp. 115-132). Mahwah, NJ: Erlbaum.

Szalma, J. L., & Taylor, G. S. (2011). Individual differences in response to automation: The five factor model of personality. *Journal of Experimental Psychology: Applied*, *17*(2), 71-96.

Szalma, J. L., Hancock, P. A., Warm, J. S., Dember, W. N., & Parsons, K. S. (2006). Training for vigilance: Using predictive power to evaluate feedback effectiveness. *Human Factors*, *48*(4), 682-692.

Szalma, J. L., Warm, J. S., Matthews, G., Dember, W. N., Weiler, E. M., Meier, A., & Eggemeier, F. T. (2004). Effects of sensory modality and task duration on performance, workload, and stress in sustained attention. *Human Factors*, *46*(2), 219-233.

Tan, K., Chan, K., & Choi, K. (2000). Detection of the QRS complex, P wave and T wave in Electrocardiogram. *First International Conference on Advances in Medical Signal and Information Processing* (pp. 41-47).

Thompson, J. (1994). Medical decision making and automation. In M. Mouloua & R. Parasuraman (Eds.), *Human Performance in Automated Systems: Current Research and Trends* (pp. 68-72). Hillsdale, NJ: Erlbaum.

U.S. Army UAS Center of Excellence. (2010). *U.S. Army Roadmap for Unmanned Aircraft Systems*. Fort Rucker, AL.

Vicente, K. J., Thornton, C. D., & Moray, N. E. (1987). Spectral analysis of sinus arrhythmia: A measure of mental effort. *Human Factors*, *29*(2), 171-182. Human Factors and Ergonomics Society.

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In M. Mouloua & R. Parasuraman (Eds.), *Automation and Human Performance: Theory and Applications* (pp. 183–200). Lawrence Erlbaum Associates.

Wickens, C. D. (1976). The effects of divided attention on information processing in manual tracking. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(1), 1-13. doi:10.1037/0096-1523.2.1.1

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and Performance VIII* (pp. 239-257). Hillsdale, N.J.: Lawrence Erlbaum. Retrieved from http://books.google.com/books?hl=en&lr=&id=oPGHfhVX6lEC&oi=fnd&pg=PA239&dq=The+structure+of+attentional+resources&ots=LBuRt1X09f&sig=4Er_zwXXep0tJXqt9Jw9hZvAJGA

Wickens, C. D. (1984). Processing resources in attention. In R Parasuraman & D. R. Davies (Eds.), *Varieties of Attention* (pp. 63-101). New York, NY: Academic Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=_KfaN4gKP8MC&oi=fnd&pg=PA3&dq=Processing+resources+in+attention&ots=S_z4Gj-4qw&sig=HDLlEIrBZ7o_RvpRaad5YbgENGc

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*(3), 449-455.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., Li, H., Santamaria, A., Sebok, A., & Sarter, N. B. (2010). Stages and levels of automation: An integrated meta-analysis. *Human Factors and Ergonomics Society Annual*

*Meeting Proceedings* (Vol. 54, pp. 389–393). San Francisco, CA: Human Factors and Ergonomics Society.

Wiener, E. L. (1984). Vigliance and Inspection. In J. S. Warm (Ed.), *Sustained Attention in Human Performance* (pp. 207-246). Chichester, UK: Wiley.

Wiener, E. L. (1985). Beyond the sterile cockpit. *Human Factors*, *27*, 75-90.

Wiener, E. L. (1988). Cockpit automation. In E. L. Wiener & D. C. Nagel (Eds.), *Human factors in aviation* (pp. 433-461). San Diego, CA: Academic Press.

Wilson, G. F. (1992). Applied use of cardiac and respiration measures: Practical considerations and precautions. *Biological Psychology*, *34*(2-3), 163-178. doi:10.1016/0301-0511(92)90014-L

Woods, D. D. (2002). Steering the reverberations of technology change on fields of practice: Laws that govern cognitive work. *Annual Meeting of the Cognitive Science Society*. Fairfax, Virginia.

Yeh, Y., & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors*, *30*(1), 111–120. Human Factors and Ergonomics Society.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*, 459-482.

Young, M. S., & Stanton, N. A. (2001). I didn't do it: Accidents of automation. In M. J. Smith, G. Salvendy, D. Harris, & R. J. Koubek (Eds.), *Usability Evaluation and Interface Design: Cognitive Engineering, Intelligent Agents and Virtual Reality* (pp. 1410-1414). Mahwah, NJ: Lawrence Erlbaum Associates.