COMPARING ATTENTION THEORIES UTILIZING STATIC AND DYNAMIC FUNCTION ALLOCATION METHODS OPERATIONALIZED WITH AN EXPERT SYSTEM

A Dissertation
Presented to
The Academic Faculty

By

Regan H. Campbell

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Engineering Psychology

Georgia Institute of Technology
June 2003

COMPARING ATTENTION THEORIES UTILIZING STATIC AND DYNAMIC FUNCTION ALLOCATION METHODS OPERATIONALIZED WITH AN EXPERT SYSTEM

Approved by:

Dr. Gregory M. Corso (Advisor)

Dr. Wendy Rogers

Dr. Bradley Fain

Dr. Bruce Walker

Dr. Dennis Folds

Date Approved: 8/20/03

ACKNOWLEDGEMENT

This paper is the result of a lengthy process, over which I have received support from many individuals. I would like to begin by thanking my family and friends, specifically my Mom and Dad, Tara, Annemarie, Aideen, and Margaret. Your constant encouragement, advice, and belief that I would finish buoyed me whenever my spirits were low. To Mark, you were my biggest champion in the early years, and I wouldn't be here without you! I would also like to thank my "kids", Bo, Odin, and Nikita, who waited patiently for me to finish working, even when they wanted to play. To everyone else who listened and had a kind word, I thank you as well.

In addition to encouragement, I received a great deal of logistical help in this process. To Eric, I can never repay you for the time you spent away from your family getting this simulation to work. Thank you from the bottom of my heart! Scott – you probably did more than anyone to ensure I was able to utilize this simulation, collect data, and make this a reality. I will never forget... To Kyle, thanks for letting me invade your life for a few weeks when everything got crazy. I owe you big, babe! I'd also like to thank those at ASI who volunteered to be my pilot participants: Mike, Rick, Brock, Mannie, Eric, and Lloyd.

Finally, I would like to thank my advisor, Dr. Corso, who helped shape my ideas and provided guidance throughout the process. I owe a debt of gratitude to you, along with the rest of my committee, Drs. Fain, Folds, Rogers, and Walker. Your insightful questions helped me refine my ideas and presentation to ensure a better final product.

Thanks to all! I am in your debt...Love, Regan

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ANOVA – Analysis of Variance

ATC – Air Traffic Control

DME – Distance Measuring Equipment

EEG - Electroencephalogram

FA - Function Allocation

Hz – Hertz

JPALS – Joint Precision Aircraft Landing System

LOA – Level of Autonomy

MANOVA – Multivariate Analysis of Variance

MART – Malleable Attentional Resources Theory

MRT – Multiple Resource Theory

PCP – Proximity Compatibility Principle

PRP – Psychological Refractory Period

SA – Situation Awareness

SAR – Synthetic Aperture Radar

SART – Situation Awareness Rating Technique

TES – Tactical Entity Simulator

TLX - Task Load Index

UCAV – Unmanned Combat Air Vehicle

WL – Workload

WWII – World War II

SUMMARY

This research examined two attention theories, Multiple Resource Theory and Neumann's view of attention, by using a secondary task methodology and a manipulation involving adaptive automation to determine if either theory was supported by the findings of this experimental study. The secondary task methodology used three tasks, with varying relationships to the primary task, to test predictions made by the attention theories. The adaptive automation manipulation involved asking participants to complete the primary task using static or dynamic function allocation methodologies. Although previous research found that dynamic task allocation should reduce attentional demands, the attention theories had different predictions about the trend of this reduction. From the adaptive automation manipulation, there was partial support for both theories, with detailed analyses showing more support for MRT. The secondary task manipulation revealed little support for either theory. In addition, there was not strong evidence showing a benefit of adaptive versus static function allocation. From these findings, task and display design recommendations were developed, as well as recommendations for improving the predictive power of the two attention theories. Also, measurement issues were highlighted, with specific recommendations given about how to improve the situation awareness and workload tools used in this study. Lastly, future research areas were suggested that could test the attention theory improvements suggested in this research, as well as utilizing the changes to the tools. This was an important question because it utilized attention theories to drive predictions about adaptive automation, whereas other research in this area simply made predictions based on previous experimental findings.

INTRODUCTION

The focus of this study is to examine two attentional theories, Multiple Resource Theory (Wickens, 1984) and Neumann's theory of attention (Neumann, 1987), to see if either is supported by the results of this experimental task. This will be done using a secondary task methodology. In addition, this paper will examine the predictions of these theories on static and dynamic function allocation methodologies. Although numerous studies have examined static and dynamic allocations, there has been little theoretical work done to examine what attention theories would predict about these two conditions. In order to do this, a specific decision aiding system will be used to support the changing allocations, so the performance and benefits of the allocation methodologies can be measured.

Attention Research

Attention is considered one of the building blocks of cognition, along with pattern recognition and memory, thus it is an important area to study. Attention is defined as cognitive resources, mental effort, or concentration devoted to a cognitive process (Galotti, 1998). Attention research began with William James, but was emphasized after WWII (Rogers, Rousseau, & Fisk, 1999). Although there are many forms of attention traditionally studied, including focused attention, divided attention, and sustained attention, this paper will only discuss divided attention, as it is the most relevant to the experimental task, using automation.

Divided Attention

According to Sanders and McCormick (1987), divided attention occurs when two or more tasks must be performed simultaneously with attention directed at both, suggesting the need for divided or multiple pools of resources. Others view divided attention as the process which attention is paid to one activity, then to another, and back to the original in rapid succession. This suggests a switching mechanism (Sanders and McCormick, 1987).

Capacity limited views of attention

Historically, most researchers have taken one of two perspectives as to how divided attention occurs: single-resource and multiple-resource theories, both of which assume there is a limit to the amount of attention that can be directed at any problem. Single-resource theories posit a pool of undifferentiated resources that are applied to all mental processes except early sensory processing (Kahneman, 1973). Thus, as more tasks are performed, the resources available decrease and performance suffers. In other words, divided attention problems occur when the task requirements exceed the resources available.

This theory would predict that dual-task performance should depend on task demands, which has been found to be incorrect in most empirical settings. According to Meyer and Kieras (1997), this led to additions to the single-resource theory by researchers such as Moray (1967), Kerr (1973), and Navon & Gopher (1979). One addition to single-resource theory is called structural interference, in which Kahneman (1973) stated that interference between dual-tasks could occur if they occupy the same mechanism of perception or response. Unfortunately, it is quite difficult to empirically separate this from a multiple resource theory, as one would have to show that central processing during multiple task performance depends on one unspecified pool of resources (Neumann, 1987), thus very little work has followed up on this perspective.

While single-resource theories are a significant step towards understanding the performance of more than one task, they have major limitations. For instance, Navon and Gopher (1979) and Wickens (1984) both commented that the manipulation of some task parameters affects two simultaneously performed tasks, while others affect just one of the tasks, suggesting more than one pool of resources. Another limitation of the single-resource theory is its inability to explain why two tasks which are both clearly attention demanding can be time-shared effectively, for example, a skilled pianist can sight-read music and engage in verbal shadowing without performance decreasing in either task (Wickens). A third issue involves the fact that resources taken away from one task cannot necessarily be allocated to the other task, as is advocated by this theory. Because of this, researchers began to theorize that multiple resources pools exist, which according to Wickens, vary according to the stage, modality, and processing code. These separate pools of resources are limited, thus shared resources will show interference with task performance (Wickens), while separate resources do not interfere with each other. These are the major tenets of Multiple Resource Theory (MRT).

Wickens (1984) elaborates on MRT with a description of the resource pools. The dimension of stage refers to the dichotomy between perception and central processing resources versus those used in selection and execution of responses. An example of this is a driver who can verbally acknowledge each change occurring in a car (a response demand) without disrupting his/her ability to maintain an accurate mental model of the environment (a perceptual-cognitive demand). Modality represents the dichotomy between resources used by the auditory system versus those used by the visual system. Two tasks presented in the same modality tend to be more distracting than presenting one task from each modality. An example of this is when an individual is able to listen to the radio and read a book at the same time, as opposed to watching two television programs simultaneously, or listening to two people talking at the same time. Lastly, processing code refers to the differences in the processing of spatial and verbal information. It is easier to process dissimilar codes simultaneously than similar codes. A secretary being able to type a letter (spatial) while simultaneously reading it (verbal) is an example of this.

Thus, as predicted by MRT, a high degree of similarity between tasks can cause confusion or interference (Wickens & Carswell, 1997). Conversely, tasks that draw on separate resources provide redundancy, as they employ separate and parallel processing. Increasing the demands of one task will be less likely to impact the other task (Wickens, et al., 1998). The ability to share resources is strongly related to the amount of workload imposed by the tasks. In particular, the mental workload of the primary task determines whether one can add a second task (Tsang & Wilson, 1997; Wickens, Gordon, & Liu, 1997).

One of the assumptions made by MRT researchers is that workload is related to the participant's performance. Although there are cases when this has not been borne out, there are also numerous cases when it has, particularly within aviation domains (e.g., see Scallen, Hancock, & Duley, 1995; Parasuraman, Sheridan, & Wickens, 2000; Raby & Wickens, 1994; Svensson & Wilson, 2002). Parasuraman and his colleagues comment that instances where a reduction in workload does not improve performance usually occur when the automation is difficult to initiate or engage or there are extensive data entry requirements. So, for well-designed automation, this should not be a factor. Scallen, et al. also note that failures to find a relationship between workload and performance may be because some researchers do not consider both components of workload. These components, the load imposed by the task and the load imposed by the individual, are both important considerations, although most researchers ignore the latter. So, the task considered to be primary (thus, emphasized by the participant) will be more likely to show the relationship between workload and performance than the secondary task(s).

Multiple resource theory is also not without its critics. For instance, Luczak (1997) and Meyer and Kieras (1997) commented that a problem with this model is it is difficult to refute (as new pools of resources can be added to explain experimental results), and it is difficult to determine the amount a specific task affects the total change in workload. In addition, Neumann (1987) commented that none of the multiple resource theories (including Wickens' theory) could account for all of the results in dual-task experiments. For instance, interference is often more specific than would be predicted by the MRT model. Conversely, interference is sometimes unspecific, where the interference does not depend on any resources being overloaded. Examples of these phenomena can be found in both motor performance and input of similar stimuli. For instance, Wickens' model posits a reservoir for movement activities, which would imply that given no other demands, the motor interference would depend only on task demands. However, there are numerous examples of specific interference, including manual/manual interference being greater than manual/vocal interference and hand/hand being greater than hand/foot (Neumann). Similarly, Wickens' model states that similar stimuli interfere more than those from a different modality. While this is often empirically supported, there are cases where this interference does not occur, as well as cases in which there is interference within one modality (Neumann; Tsang & Wilson, 1997). For instance, in the auditory modality, Treisman and Davies (1973) found that two tones interfere with each other more than a tone and a word. Similarly, in the visual modality, two stimuli interfere with each other more if they are in different objects or different dimensions are attended to (Neumann). An example of less specific interference than would be expected is found in dual-task response time experiments, where the number of stimulus-response pairings are more important than the type or modality of the tasks (Neumann). Allport (1989) describes alternative explanations for many of the common capacity-limited findings.

Neumann's View of Attention

Because of these problems, some researchers have begun to look at attention in new ways. Up to this point, attention has been considered a capacity limited system. In other words, attention exists because the brain is not capable of handling multiple inputs at the same time. However, there is no established limit on the information that the brain can pick up at once (Neumann, 1987; Meyer & Kieras, 1997), possibly making this assumption an invalid one. A different way of viewing attention is espoused by Neumann, who views the bottleneck to be in the effectors of activity, rather than in the brain. Thus, attention is a means to prevent multiple simultaneous actions. In support of this, Meyer and Kieras commented that central processing bottlenecks are actually occurring because instructions constrained the participant's responses. They found unconstrained participants had flexible strategies for scheduling processes to satisfy task requirements, thus the apparent bottleneck disappeared.

This view of attention as an effector-limited system began with Neisser and Allport in the 1970s. Neisser (1976) commented that performance limitations were due to physical or coordination problems. For instance, interference in motor outputs may occur because it may not be physically possible to complete two actions at the same time (e.g., type and juggle at the same time). Similarly, two visual inputs may interfere with each other, if one masks the other. Coordination problems may occur if two tasks were learned separately and then combined, requiring further learning. This can be particularly true in emergency situations, as the combination of a task and the emergency task was never practiced. Allport (1980) elaborated these findings by saying that all interference is data-specific (cross-talk causing the wrong action for an input) or function specific (competition for the same action).

Neumann (1987) commented that although all of these phenomena are true, they rarely occur in a laboratory, thereby failing to explain many empirical results. For instance, experimenters do not create tasks in which the movement activities conflict with each other. In addition, humans do not spontaneously have problems where they try to carry out conflicting activities. Because of this, Neumann's theory posits that a function of attention may be to prevent these types of interference. In other words, the capacity limitation is not a processing limitation, but rather a lack of effectors to carry out actions and inability to specify the parameters needed to carry out all actions at the same time. So, attention is related to the control of actions, as we must combine various systems (e.g., visual attention, motor activity) in order to coherently carry out an action (Allport, 1989). It provides a means for the human to select which action is the critical one to carry out at that time.

Neumann (1987) examined empirical findings in light of his view of attention. With respect to the limitation of effectors, he posits that effectors have a blocking mechanism that is used until a task is completed. This is true unless the action is interrupted, which requires an orienting mechanism to explain the interrupting of actions. This notional orienting function is supported by the research into preattentive processing, as this stage provides the ability to attend to unexpected, but important new stimuli. When considering preattentive processing as an orienting function only, one finds support in both the early and late selection empirical results. This interrupting mechanism is also supported by research into the Psychological Refractory Period (PRP), as the "refractory period" is actually the interrupting of one action and the triggering of a new one. If an action is not completely interrupted, one can experience crosstalk between actions.

A second area Neumann discusses is action planning. He commented that multiple actions are possible, as long as they are coordinated through action planning. Action planning allows one to perform concurrent actions if the tasks are coordinated or scheduled using different resources. If the actions conflict, one or both of the actions must be modified, or the most important action must be selected for action (Allport, 1989). Practice helps in the specification of actions, as well as the planning of their coordination, making it easier for them to be integrated (Neumann, 1987). Action planning is made more demanding if one tries to add a new task or adds to the condition action rules, like requiring two actions instead of one, for existing tasks. It is also made more difficult by increasing the problems of coordinating two tasks that have the same effectors. In other words, the demands of the dual task will increase if the number and/or required steps of the condition-action rules increase or if the integration between the tasks is more demanding.

A third major tenet of Neumann's view of attention is that attentional selection (focusing) is based on the type of object. Thus, whether or not information can be attended to simultaneously depends not on information load, but rather on structural factors and spatial location of the information. This is supported by findings such as the Stroop effect and findings where form and color can be processed simultaneously for an object (Neumann, 1987). In addition, Allport (1989) stated that secondary tasks that are far away result in worse performance than secondary tasks that are close because information that is close could be simultaneously attended to, whereas far information requires the operator to switch back and forth between the locations.

The third tenet is in contrast to the Proximity Compatibility Principle (PCP), which predicts that integrated tasks will benefit more from integrated displays than tasks that require focused attention (Haskell & Wickens, 1993). However, if the tasks are not integrated, they will be hindered by an integrated display, with more benefit from a separable display. Separable displays have a lack of interaction among dimensions, and allow individuals to selectively attend to individual dimensions (Coury & Boulette, 1992). In this study, the two tasks are low compatibility, as they are dual tasks performed concurrently that do not have interaction between the information sources (i.e., metric similarity, functional similarity) or similar computational routines (see Wickens & Carswell, 1997). Thus, they should benefit from a separable display, where the components are noticeably separated on the screen, as items that are not in close proximity cannot be processed together (Wickens & Carswell). The correct display should result in faster, more accurate performance, as well as lower workload (Haskell & Wickens). To summarize, PCP states that far information should result in better performance because the tasks are separated, whereas Neumann's theory states that the close information should result in better performance.

Although Neumann's view of attention appears to be well thought out, there has been very little research that has examined it empirically. This may be because it is a vastly different view of attention or because there is a large body of evidence in support of Multiple Resource Theory. Regardless, the current research will attempt to determine whether either MRT or Neumann's views appear to be supported by the results of this experimental task. This will be done using a secondary task methodology, described further in the methods section, which compared the two theories. In general, there are three tasks that have been selected to help differentiate between the two theories of attention. Two of the secondary tasks were developed to see if MRT is supported: one similar to the primary task and one different. If the similar secondary task results in worse performance (in terms of response time and percent correct of responding to the task) than the dissimilar task, Multiple Resource Theory would appear to be supported.

The third task is designed to elucidate the feasibility of Neumann's view of attention. It is the same as the similar task mentioned above, however it will be in the periphery (as such it should result in worse performance than the dissimilar task if MRT is correct because it taxes the same resources as the primary task). This task, in combination with the similar task from above, is used because Neumann's theory states that information could be attended to simultaneously if it is located close to the major information. Therefore, this theory would predict that the similar task in the foveal area should always result in better performance (response time and percent correct) than the periphery task, regardless of the number of targets per minute. It is not clear how the dissimilar task should be affected, but it seems as if it would fall between the two similar tasks. In other words, it should not be affected by number of targets per minute and the performance should be worse than the similar task in the foveal area, but better than the similar task in the periphery. So, to summarize, MRT would predict an ordering of dissimilar task, then both similar tasks, whereas Neumann's view would predict similar task in the foveal area, dissimilar task, and then similar task in the periphery of the screen.

The predictions of Proximity Compatibility Principle are that the dissimilar and similar task in the periphery should be significantly better than the similar task in the foveal area, as both the dissimilar and similar task in the periphery are separated from the major information, therefore not processed at the same time as the primary task information. These predictions are similar to MRT predictions, which are ordered dissimilar, then the two similar tasks. So in the case of MRT, the two similar tasks are not able to be distinguished, whereas PCP cannot distinguish dissimilar and similar in the periphery.

Attention Theories and Automated Errors

After determining which theory appears to be supported with the secondary task methodology, design recommendations can be made to reduce the divided attention errors made in the experimental environment. MRT would predict that errors would occur anytime workload is exceeded or anytime the dual-tasks interfere with each other. Neumann's view, on the other hand, would predict errors would occur anytime participants are not oriented to the correct information or they cannot specify the necessary parameters to carry out a task. Because of these different predictions, design recommendations based on these theories would be quite distinct.

There are two errors of divided attention that are commonly discussed in automation research: a loss of situation awareness and a loss of mode awareness. (For a more detailed discussion of automation problems, see Campbell, 2001). While both theories could explain how these errors occur (through either excessive workload or incorrect orientation), the recommendations on how to reduce the errors are different.

Situation awareness. Situation Awareness (SA) is "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley, 1996, p. 163). As such, situation awareness can include such things as awareness of direction and speed of travel, awareness of the number of events coming up, and time since last failure. Although SA is critical for effective decision making and human performance in a dynamic environment (Endsley & Kiris, 1995), it is considered separate from decision making because high SA does not guarantee good decisions and vice versa. For instance, a person can make the wrong decisions if they have incomplete SA. Conversely, they can have a complete understanding of the situation and make a bad decision because they are inexperienced (Endsley, Farley, Jones, Midkiff, & Hansman, 1998).

There are many problems that can occur if a participant does not have good SA. Collectively, these problems are known as out-of-the-loop performance problems. These problems occur because of difficulty understanding the system (due to the inherent properties of the system, poor interface design, and inadequate training), moving from an active processor to a passive recipient of information, and changing the type of feedback provided to operators (Endsley, 1996; Endsley & Kiris). In one study, Endsley and Kiris found that the most important component in the loss of SA appeared to be the lack of active processing. Active processing likely makes the task more interesting and reduces boredom, making it easier to attend to it. This points to a clear benefit of allowing participants to control the system whenever possible (e.g., when the attentional demands are lowered).

On the opposite side of the scale, when tasks are quite difficult, some efforts to improve SA are actually directed at methods to control excessive workload. If the overall level of workload becomes excessive, then the reduction in performance can occur in SA, decision making, or actions (Endsley, 1993), thus reducing workload is helpful. This may be because more attentional resources are available to focus on the task. Collectively, the findings point to an advantage of being able to change task allocations dynamically because SA is improved with reduced workload at high stress times and with increased control at low stress times.

Interestingly, this discussion of SA research points to support for both attention theories discussed. The first, MRT, is supported by the research that suggests that reducing workload would be helpful. MRT would predict that automation problems occur because too much attention and workload is directed at a different task (other than maintaining SA). Thus, a reduction of workload or attentional resources on other tasks should allow resources to be used for SA maintenance. On the other hand, the research about active processing would seem to support Neumann's view of attention, as active processing would suggest a participant was focused on the correct information therefore able to maintain their SA. If they were not focused on the correct information, they would lose SA and have difficulty reorienting themselves.

Mode awareness. Mode awareness problems occur after mode transitions, when the participant is not aware there has been a change, thus it is similar to a lack of SA. It can result in errors in setting up the automation (Wiener & Curry, 1980) or incorrectly interpreting feedback from the automation. Mode awareness issues occur because there is a proliferation of modes in complex systems, resulting in an increased risk of inadvertent activation of modes by the operators (Sarter & Woods, 1995). This creates a cognitive demand on operators, in which they must know how the system works in each different mode and how to manage each new set of options in different contexts (Sarter &

Woods). In other words, operators must know how the functions work, which mode to use when, how to switch between modes, and which mode is currently active, which is a huge memory and attentional demand. Javaux (1998a) found that pilots were not always able to assess which modes were active, what the modes were doing, and what the modes were going to do next. In short, they had an obvious lack of understanding of the automation, although they were highly trained (Javaux, 1998b; Javaux & De Keyser, 1998).

There are two recommendations for improving mode awareness. The first is to provide the participant with information about the mode change triggers or provide them with adequate feedback to make sure they notice the change (Degani, Shafto, & Kirlik, 1999). The other method to improve mode awareness is to reduce the number of modes (Sarter & Woods, 1995), thus allowing resources to be used for attending to the current mode and situation.

Similar to SA, this discussion of mode awareness research can be explained by both theories, with one recommendation for supporting mode awareness pointing to MRT research and the other pointing to Neumann's view. MRT suggests that reducing the attentional demands of determining what mode is active would reduce the workload and improve mode awareness. This reduction could come via a reduced number of modes. Neumann's view of attention would suggest that mode awareness problems occur because a participant is focused on different information than the mode information, preventing them from attending to the mode change. Thus, the ability to highlight the mode change through adequate feedback should help the participant to attend to the change.

Summary

Much attention research has focused on Multiple Resource Theory, which has proven to be an important theory in the study of attention. This research will examine the various pools of attention and made predictions about the types of interference experienced by participants in a dual-task environment. Specifically, MRT claims that similar types of information will interfere with each other, whereas different types of information will not interfere. In order to determine the similarity between tasks, they are evaluated on scales of mode, stage, and processing code. Although MRT has explained many empirical results, it has critics that point to examples that cannot be explained. This has led a group of researchers to view attention as an effector-limited problem, rather than a capacity-limited need. This theory, championed by Neumann, claims that attention is used to prevent simultaneous actions or prevent actions that have not be properly specified from occurring. Thus, whether multiple actions can be attended to is a function of the type of information (spatial location and structural factors), rather than the information load. These theories, with their differing predictions, will be compared through a secondary task methodology to determine if either is supported by the empirical results. In addition, design recommendations based on the supported theory will be discussed which are aimed at alleviating some of the divided attention problems experienced by participants in this environment.

Adaptive automation

Clearly divided attention research has implications in a number of domains, including the use of automation. One way that has been suggested to improve operator performance in this environment is to use adaptive automation (adaptive function allocation) to reduce attentional overloads and workload. This type of function allocation is designed to provide increased interaction in low workload times and reduced interaction in high workload times. In theory, this switching of levels of automation should allow the participant to be more involved in the system, while stabilizing attention and workload. It should also allow active processing, particularly at low workload times, helping participants to build Situation Awareness (SA) and mental models.

Traditional function allocation

Traditional or static Functional Allocation (FA) is a technique that is used to assess the capabilities of humans and machines, and allocate tasks based on these capabilities (Scallen & Hancock, 2001). There is often confusion associated with the term function allocation, as it is often used synonymously with the term task allocation. Task allocation is a decision about which entity should perform a task, given the current capabilities of each component (Hancock & Scallen, 1998). Tasks are steps that enact strategies and achieve goals. Therefore, they are discrete events that have a beginning and an end with some transformation of the world state (Hancock & Scallen). According to Hancock & Scallen, functions are aggregates between tasks and resources, representing goals, strategies, and tasks. Function allocation is then defined as the assignment of tasks and resources to the human or the machine. Thus, function allocation is a broader term that includes both task allocation and resource allocation, which is the process where energy is mobilized towards a task (Hancock & Scallen).

Traditional FA is a static allocation of functions to the human or the machine. This technique was first devised by Fitts (1951), who developed a list of capabilities, which has been used quite extensively. The Fitts List was originally designed using the information processing paradigm, as a way to stimulate basic research to identify the abilities of humans and machines. Unfortunately, this has led to terminology and concepts that favor the machine (Hancock & Scallen, 1998). Further, the Fitts List has been interpreted as a basis for design recommendations, which has led to its failure in a number of static allocation tasks. This is not surprising given the capabilities of machines, as well as training practices, have changed significantly since 1951 (Hancock & Scallen).

Traditional FA has been criticized for a number of reasons. First, the technique is static and acontextual (Scallen & Hancock, 2001; Scallen, Hancock, & Duley, 1995), so it is insensitive to the environmental conditions or how they change during task performance. A second problem is traditional FA fosters comparison and division of responsibility, rather than cooperation (Scallen & Hancock). This division of responsibility is appropriate when only one component can perform the task (Kantowitz & Sorkin, 1987), but not appropriate when both can perform the task. Both of these problems can lead to a third issue: extended periods of human or machine control. Extended human control can lead to increased fatigue, increased workload, and reduced effectiveness, whereas extended machine control can lead to vigilance problems (Scallen & Hancock).

However, static FA could have advantages because it provides active processing which should build SA and mental models. A second possible advantage to static FA is that the stimulus-response pairing is always consistently mapped, which has been shown by numerous researchers to be superior to varied mappings (Schneider & Fisk, 1982; Strayer & Kramer, 1994). In addition, static FA would eliminate mode awareness problems where a participant is not aware they are in charge of a task.

Adaptive function allocation

To combat the problems associated with traditional FA, adaptive function allocation was developed. It can circumvent many of the problems with automated systems by modifying the level of operation in response to changing demands placed on the operator (Scerbo, Freeman, & Mikulka, 2000). It does this by controlling the allocation of functions between the human and the machine, based on environmental factors, operator competence, or psychophysiological factors (Haas & Hettinger, 2001; Scallen & Hancock, 2001). This shifting is only done when human performance in the system needs support to meet operational requirements; without that, humans should be in control (Andes, 1990; Bennett, Cress, Hettinger, Stautberg, & Haas, 2001). However, an assumption in dynamic function allocation is that both the human and the machine are capable of performing a task. Responsibility for the task is given to the decision maker who has resources available.

Adaptive automation can create a tighter coupling between the participant's workload and the degree of automation (Parasuraman, Bahri, Deaton, Morrison, & Barnes, 1992), as the level of automation changes based on the workload. There are four advantages to dynamic task allocation: (Greenstein & Revesman, 1986; Hancock & Scallen, 1998; Tattersall & Morgan, 1997)

1) System resources are used more fully
2) Operator workload is moderated and remains more constant
3) It enables humans to have a more flexible role – increasing SA, increasing job satisfaction, and decreasing skill decay
4) It improves system performance and is more acceptable to operators

These benefits may not, however, outweigh the mode awareness problems that could occur when the system is continuously switching between modes.

Three methods are generally used for triggering a change in the level of automation. The first is to monitor the participant's performance and make a change based on a criterion of performance. This method is limited because it requires continuous monitoring of the system in order to be effective. The second method is to use a performance model, which makes a change based on the future state of the system, given the current state, external events, and participant performance. This method also requires monitoring of participant performance. Lastly, one can continuously monitor psychophysiological (e.g., pupil diameter, eye blinks, electroencephalogram (EEG), heart rate, and respiration) measures to determine when a change in function allocations should occur. While this system is advantageous because it does not require interpretation of overt behavior, it is problematic because it is expensive to obtain and susceptible to "noise", making interpretation difficult (Scerbo, Freeman, & Mikulka, 2000).

Experimenters have used these three methods of changing function allocations successfully, although all need further refinement. Because of equipment constraints, this experiment will not attempt to further examine psychophysiological function allocations. Of the other two methods, this study will specifically concentrate on the performance-based adapting because it is easily implemented and its operationalization is the most clearly defined. In addition, the ability to continuously monitor, which is one of the drawbacks of this approach, is not an issue in the system that will be used in this experiment.

Level of autonomy

An important issue in adaptive function allocation is the level to which the switch is made. Many researchers have examined changes in function allocation between fully manual (where the participant did everything) and fully autonomous (where the machine did everything), which can be quite a dramatic change for a participant. However, there are clearly a number of levels between these end points, which may provide less disruptive changes in allocation. The first taxonomy of automation levels was published by Sheridan and Verplank (1978) (see Figure 1), with more recent taxonomies adding little to the original work (Moray, Inagaki, & Itoh, 2000). In the Sheridan and Verplank classification scheme, levels 1-4 show humans having responsibility for decision making, levels 8-10 show autonomous behavior, and levels 5-7 showing the collaboration between humans and machines. Many have proposed that the level should change dynamically depending on the situation (Moray et al., 2000), as opposed to switching between only the extremes.

1. The human does the whole job of planning, selecting options, and preparing the machine, up to the point of switching on the machine to carry out the actions.
2. The human asks the computer to suggest options and selects from those options.
3. The computer spontaneously suggests options to humans.
4. The computer suggests options and proposes one for the human to use.
5. The computer selects an action and performs it if the human signals approval.
6. The computer selects an action and performs it unless the human intervenes to cancel it.
7. The computer chooses an action, performs it, and informs the human.
8. The computer makes and implements decisions and informs the human only if asked to do so.
9. The computer makes and implements decisions and informs the human only if it feels it is warranted.
10. The computer does the entire task autonomously.

Figure 1. Sheridan-Verplank Scale of Human-Machine Function Allocation in Automated Systems (Sheridan & Verplank, 1978)

Relatively few experiments have examined dynamic allocations using multiple levels of autonomy. Moray, et al. (2000) recently used three levels in an experiment. The decision to switch levels was based on the time-criticality of the failure, with level 4 used when control is left to the human, level 6 when there is time for human intervention, and level 7 when the case is extremely time critical. Their task was to control the temperature of a water plant while monitoring and treating pipe faults (breaks and leaks). An interesting caveat in their experiment was that even at level 7, the computer action might be wrong, so the human was required to monitor and take action if there was a misdiagnosis. Moray, et al. found that utilizing adaptive automation resulted in a significant decrease in the likelihood of an accident and a reduction in the amount of fluid leaked.

Rudolph (2000) modeled five levels of autonomy: full manual, management by delegation, management by consent, management by exception, and automatic. In full manual conditions, the participant was responsible for performing all tasks. In management by delegation, the participant was able to delegate the control of the resource management task. Similarly, management by consent and management by exception required consent of the participant or disapproval of the action, respectively. Lastly, automatic mode occurred when the system operated the resource management task without participant involvement. He found significant differences only between full manual and full automatic, with fully autonomous operations resulting in better detection of events. This was counter to his prediction that fully autonomous operation would result in peripheralization. Clearly, more research is needed that combines levels of autonomy and adaptive allocation, to determine whether multiple levels help operator performance.

This research will use three levels of autonomy for the dynamic automation, which correspond to the areas of collaboration between the human and machine on Sheridan and Verplank's Scale. They are the permission (level 5), veto (level 6), and autonomous modes (level 7). Participants interacting with the static function allocation will experience just the veto mode.

Summary

In general, adaptive function allocation has been found to improve system performance by helping participants maintain involvement with the system and reduce attention problems, but it may increase mode awareness problems. On the other hand, traditional (static) allocations may result in operator out-of-the-loop performance, but no mode awareness problems. The benefits of the two methodologies will be further explored with this experimental task.

Attention Theories and Adaptive Automation

As is apparent in the last section, many researchers believe that adaptive automation is superior to static automation. They can point to variables that show improved performance (e.g., increased SA) and posit that this suggests a reduction in attentional demands, but there is no discussion of attention theories to support this. In fact, an examination of the two theories discussed in this paper indicates two differing predictions about the benefits of adaptive automation relative to static allocation.

To examine these different predictions, both the primary and secondary tasks will be examined. For the primary task, response time and percent correct will be computed using responses from when the operator is performing in the permission and veto modes (autonomous mode will be considered missing data in the data analyses), as permission and veto modes represent the operator's performance on this task. For the secondary task, the operator performs the task during all three modes, so response time and percent correct will be measured throughout task performance. If the operator does not respond to the secondary task, it will be considered incorrect. Participants will be instructed to weight the primary task more than the secondary task.

MRT researchers believe that the ability to share resources is strongly related to the amount of workload imposed by the tasks. Thus, for the primary task, these researchers would predict that as the number of targets per minute increases, the reduction of workload inherent in the dynamic allocations should offset the increasing demands of the task, whereas the static mode would not have a similar counteracting of the effects of increasing demands. So, for the dynamic condition, because workload should be relatively constant throughout (as increasing task demands are offset by less workload as the level of autonomy increases), performance should also not be significantly different for any number of targets per minute. With the lower number of targets, when the participant is in the dynamic mode and interacting with the system at a lower level of autonomy (mostly permission mode), the dynamic condition is expected to result in higher workload than the static condition, which is in the veto mode. As the number of targets per minute increases, participants in the dynamic mode will be mostly operating in the veto mode, so operator performance on the primary task is expected to approach static performance (see Figures 2 and 3).

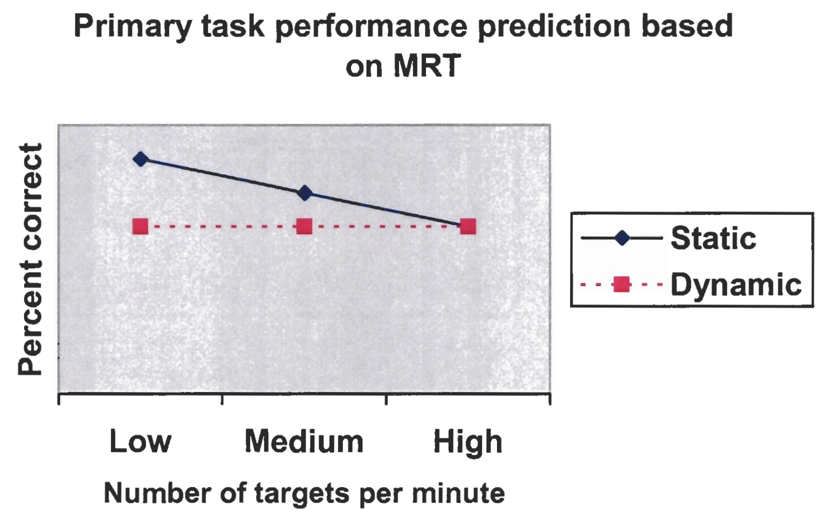

Figure 2. Predicted percent correct by number of targets per minute based on MRT

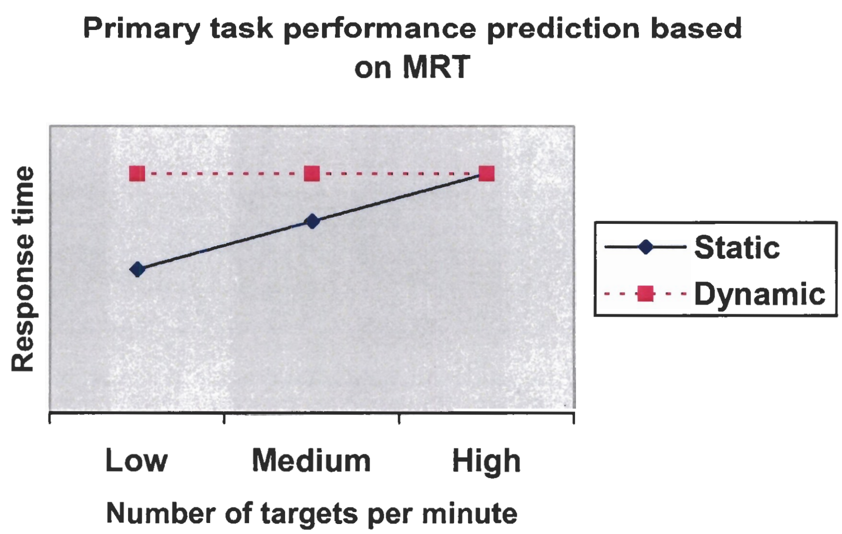

Figure 3. Predicted response time by number of targets per minute based on MRT

For the secondary task performance, MRT would predict, as many researchers do, that the dynamic condition would result in better performance than static allocations at high levels of workload. Again, this prediction is derived from the fact that MRT researchers believe that the ability to share resources is strongly related to the amount of workload imposed by the tasks. Participants in the dynamic mode are expected to have stable performance on the secondary task because as performance on the secondary task declines, the primary task is bumped to a higher level of autonomy. This higher level of autonomy will reduce the demands on the primary task and participants will have more available resources to concentrate on the secondary task. Thus, the changing levels of autonomy will offset the increase in demands on the dual task. The static mode will result in declining performance on the dual task as the number of targets per minutes increases, which will show declining performance in the secondary task. In the lower number of targets, when the participant is interacting with the system at a lower level of autonomy (mostly permission mode) for the primary task, the participants in the dynamic condition have higher workload for the dual task than the static condition, which is in the veto mode. This should result in worse secondary task performance in the dynamic mode than the static mode. However, at higher numbers of targets per minute, the dynamic mode will result in primary performance in the autonomous mode for a larger percentage of the time, so there will be less workload for each target (and more targets). This will maintain performance on the secondary task at a steady value, while static performance declines steadily. To summarize, performance in the static allocation condition is predicted to be significantly better at low number of targets per minute and decreasing as number of targets per minute increases. Performance in dynamic allocation condition is predicted to be significantly worse than static at low number of targets per minute and significantly better than static at high number of targets per minute (see Figures 4 and 5).

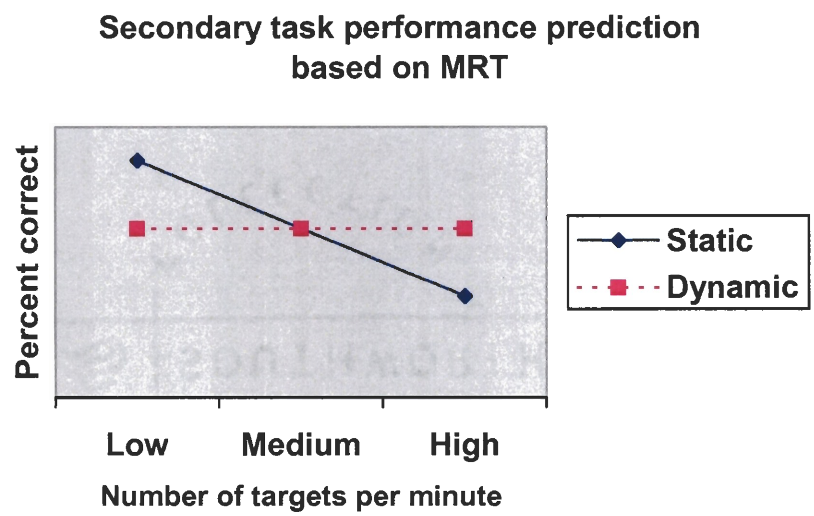

Figure 4. Secondary task percent correct predictions based on MRT

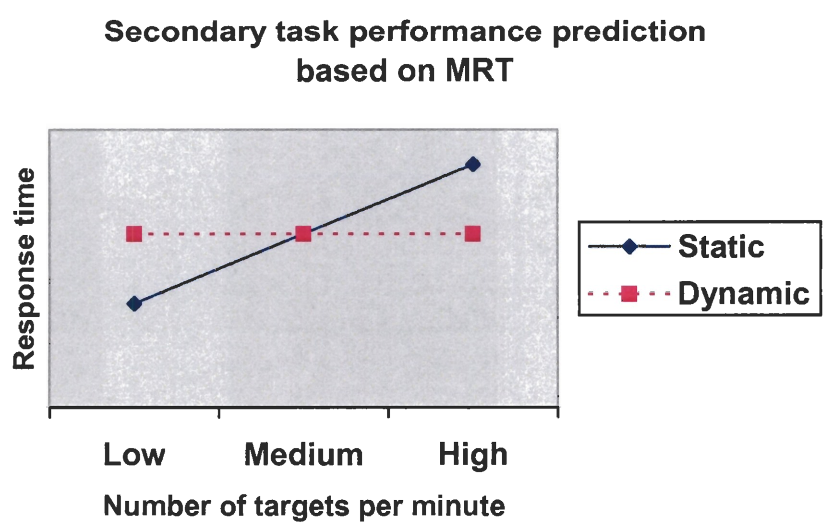

Figure 5. Secondary task response time predictions based on MRT

The predictions above for MRT assume that the increase in the resources available in the dynamic condition will be divided equally between the primary and secondary task to offset the increased demands of having more targets per minute. However, it is possible that the primary task could be focused on, leading to improvements in performance on this task. If this occurs, there will be a decline in performance on the secondary task. Thus, for the primary task, the participants in the dynamic condition will show an increase in performance, while the participants in the static condition are expected to show a decline. The secondary task, on the other hand, will show both the static and dynamic conditions with a decline in performance. This result was not borne out in the data, thus participants appeared to divide extra resources evenly between the two tasks.

In addition to the predictions above, MRT would also predict an interaction between performance as the number of targets per minute increased and the secondary task. In the case of MRT, the prediction is that the similar task will have a larger difference in performance than the dissimilar task when comparing the number of targets per minute, as the workload will change more for the similar task. This occurs because the similar task is predicted to be more distracting to the operator; with this distraction becoming compounded as the number of targets increases. The dissimilar task is expected to result in a lower level of distraction at the lower number of targets, with less of an increase as the number of targets increases.

Neumann's theory has different predictions. To review, he states that multiple actions are possible, as long as they are coordinated through action planning. Action planning allows one to perform concurrent actions if the tasks are coordinated or scheduled using different resources. If the actions conflict, the most important action must be selected for action (Allport, 1989), which in this case is the primary task. Action planning is made more demanding if one tries to add a new task or adds to the condition action rules. It is also made more difficult by increasing the problems of coordinating two tasks that use the same effectors. For the primary task, the condition action rules will not change in the static condition (as they are always in veto mode) and they will be told to concentrate on this task, so demands should not increase and performance should not decline (i.e., performance should not be significantly different at any of the number of targets per minute). In the dynamic mode, the condition action rules get easier as the number of targets per minute increases, so there is expected to be a significant improvement in performance. This is predicted to occur because a larger percentage of time is spent at the higher level of autonomy (veto) where the rules are easier. Because veto mode is the mode used in the static condition, performance is expected approach static performance. To summarize, performance in the static allocations is predicted to show no change throughout the task, whereas the performance in the dynamic condition is expected to show an improvement with increasing number of targets per minute because action planning gets simpler (see Figures 6 and 7).

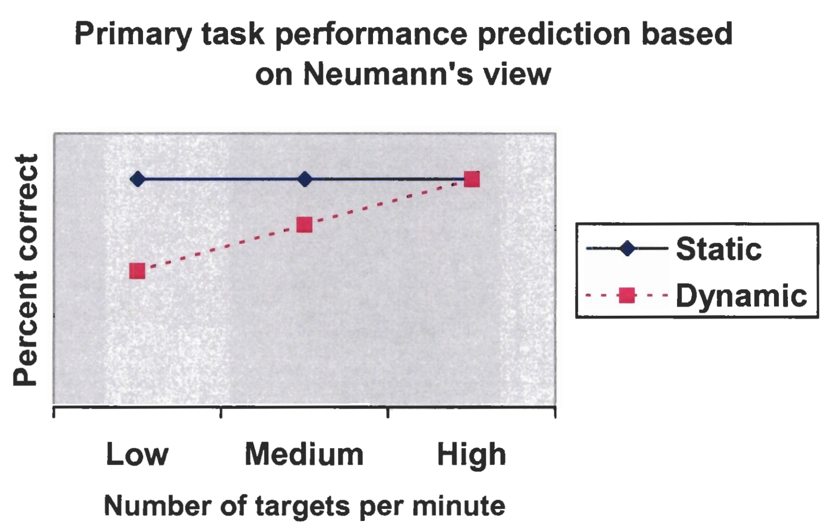

Figure 6. Predicted percent correct on the primary task by the number of targets per minute based on Neumann's theory

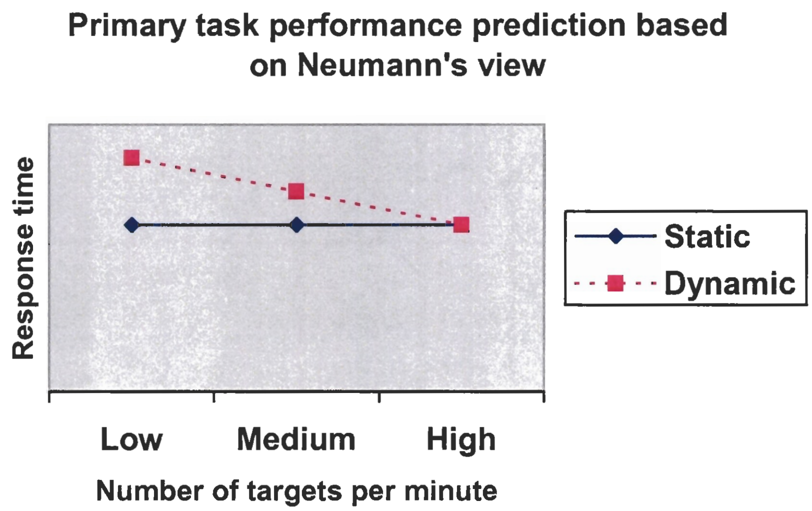

Figure 7. Predicted response time on the primary task by the number of targets per minute based on Neumann's theory

Under Neumann's view, the secondary task performance is predicted to be similar to MRT's predictions, although for different reasons. Because the primary task should be emphasized in the task performance, the secondary task is expected to suffer as the number of targets per minute increases. For the dynamic mode, this should lead to an increase in the level of autonomy on the primary task as the number of targets increases. Because there were fewer actions that needed to be planned for each target (with more targets), the dynamic mode should show relatively stable performance as the number of targets increases. Participants in the static mode for the secondary task are expected to show a decline in performance as the number of targets per minute increase. This is predicted because neither the primary or the secondary task are getting any easier, so the action planning becomes increasingly difficult as the number of targets per minute increases. At the lowest number of targets, the permission mode is most prevalent in the primary task for the dynamic mode, making action planning easier for participants in the static condition. At the highest number of targets, the autonomous mode takes over more of the primary task in the dynamic condition, allowing reduced interaction with the system and improved action planning for the secondary task to compensate for the increased number of targets to be scheduled. Thus, the prediction is for participants to experience a decline in the static mode and stable performance in the dynamic condition (see Figures 8 and 9).

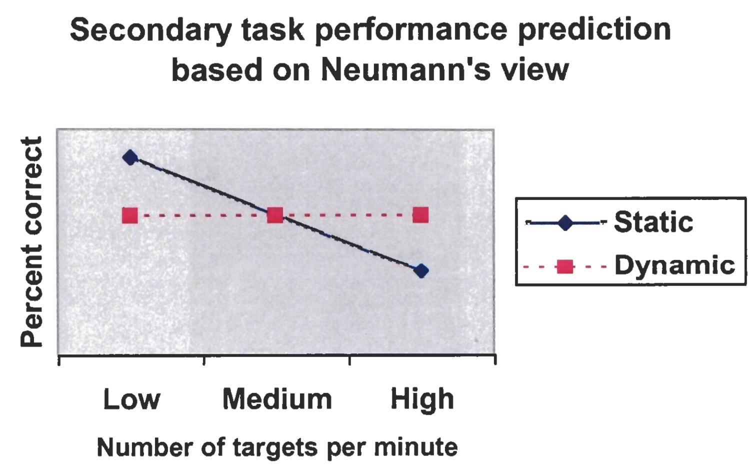

Figure 8. Secondary task percent correct predictions based on Neumann's theory

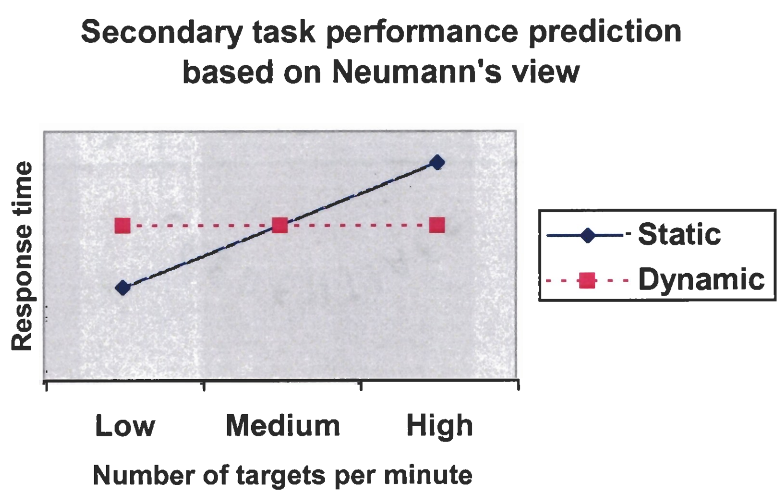

Figure 9. Secondary task response time predictions based on Neumann's theory

Neumann's view does not predict an interaction between performance in the secondary tasks as the number of targets per minute increases, as these tasks do not affect the condition action rules or the action planning. Although the similar task in the periphery area is predicted to result in worse performance than the similar task in the foveal area, the pattern of performance is not predicted to be different as the number of targets per minute increases.

Thus, the attention theories discussed in this paper have quite different views on how dynamic and static function allocations behave. The trend information is opposite, with one theory predicting no change in performance on the dynamic condition and the other predicting no change in performance on the static one on the primary task performance measures. Thus, this manipulation and the subsequent measurement of performance should provide additional insight into the two attention theories.

Automated Domains

The experimental task, as mentioned previously, will focus on the use of automation to examine both the attention theories and function allocation. Automation is defined in this paper as a system that accomplishes (partially or fully) a function that could be done by a human or that was once done by a human. Automation was chosen for two reasons, 1) it places new cognitive demands on participants, and 2) many of the problems associated with automation are attention based.

With regards to the new cognitive demands, Hollnagel and Woods (1999) found there has been a shift from requiring operators to use perceptual motor skills for control to using more cognitive activities, like decision making. This has increased the mental workload and resulted in additional burdens involving new tasks, new cognitive demands (including new management tasks, new attentional demands, new coordination and communication tasks, and new knowledge requirements), role changes, and more data (Woods, 1996). For instance, when operators use automation, they have to develop new mental models about how to interact with the system. This can result in additional task and cognitive demands, as one tries to learn the operation of the automated system, which is often constructed to have little feedback. As a result of the changes introduced by automation, it is more difficult for operators to develop accurate mental models and attend to changing conditions. This necessitates a system that can support the building of mental models and situation awareness, as well as helping operators to maintain attention on the task.

The second reason for choosing an automated environment was that divided attention problems often occur in this domain. A select group of these problems was discussed earlier in the paper. Specifically, there were mode and situational awareness problems. These problems highlight the predictions made by the two attentional theories and point to specific design recommendations that follow if a given theory is supported. Thus, after exercising the two manipulations in this experiment, design recommendations based on the theory that is supported can be discussed to alleviate automation problems.

Summary of the Major Questions

To summarize the major questions of this experiment:

1) Is MRT or Neumann's view of attention supported by the empirical results?

A secondary task paradigm will be employed to determine which (if either) of these views of attention appears to be supported by the performance on the secondary tasks. This will be done to further elucidate the attention research, as well as potentially provide design recommendations. The two attention theories have differing predictions about the ordering of the secondary tasks. For MRT, the dissimilar task is expected to result in faster response time and more accurate percent correct performance than either of the similar tasks. Neumann's theory predicts that the similar task in the foveal area will result in faster and more accurate performance than the similar task in the periphery. The dissimilar task is predicted to fall somewhere in between the two similar tasks.

2) Are either of the two attentional theories supported by the dynamic versus static automation manipulation?

As mentioned, the two attentional theories have different predictions as to the primary task performance trends expected in the dynamic and static function allocation conditions. For primary task performance, MRT predicts that performance in the dynamic condition should result in stable response time and percent correct as the number of targets per minute increases. Participants in the static condition are predicted to show slower response times and less accurate percent correct performance as the number of targets per minute increases. Neumann's theory, on the other hand, predicts stable primary task performance for participants in the static condition. Participants in the dynamic condition are predicted to show faster response times and more accurate percent correct performance as the number of targets per minute increase. Both of the theories predict that the secondary task performance will be stable in the dynamic condition and slower and less accurate in the static condition, as the number of targets per minute increase. This manipulation is designed to provide additional insight into the attention theories, as well as providing theoretical support for adaptive automation studies.

3) Is there a difference between performance in the dynamic function allocation versus the static function allocation?

This will be tested using an ANOVA that compares performance in the static allocation condition versus the performance in the dynamic condition. The hypothesis was that participants in the dynamic FA should show better performance than participants in the static FA because dynamic allows increased interaction at low stress times and less interaction at high stress times, thereby stabilizing the participant workload.

3a) In what ways is it beneficial? Workload reduction, increased SA, system performance?

This will be tested using an ANOVA with these items being the dependent variables in the task. Workload will be measured for both static and dynamic function allocations using the NASA TLX, with the prediction being the performance in dynamic allocation will produce lower workload assessments than performance in the static allocations. Similarly, situation awareness is predicted to be better for participants in the dynamic condition, as they have more resources for SA as the task gets difficult and more involvement as the task gets easier, leading to better situation awareness. The SA and workload measures will be global measures (i.e., they do not provide a measure for each level of the number of targets per minute) because the focus is on the theoretical questions in the comparison between levels of adapting rather than differences in the levels of the number of targets per minute, and this prevents having to interrupt the scenarios to get measurements.

Lastly, operator performance in terms of response time and percent correct of secondary task performance is predicted to be better for dynamic allocations, as the system is expected to adjust based on participant performance on a secondary task, removing primary tasks as demands increase. Although this condition will not be completely equated between the static and dynamic modes because at the highest level of autonomy the computer is doing tasks for the participant, it should be comparable because at the lowest level of autonomy, the participant had to do additional tasks. In addition, the more important use for the performance data is in comparing the two attention theories, rather than the allocation methodologies.

4) Does dynamic function allocation reduce automation errors?

The automation errors mentioned in the introduction will be evaluated to determine if they are reduced by the use of dynamic task allocation, as is claimed by proponents of adaptive automation. The first is the inherent reduction in situation awareness that comes from using an automated system. Dynamic function allocation is predicted to improve the SA of the participant relative to a static allocation. This will be directly measured using the SART tool, described later.

Another error that will be indirectly assessed is mode awareness. Because there are no mode changes in the static task allocation condition, there could be no comparison in this task. However, the debriefing questionnaire will ask participants whether they know what mode they ended in, as well as if they are satisfied with the notification of mode changes. This will allow an evaluation of whether the display provides sufficient mode awareness to the participants. The display is designed to provide mode awareness by notifying the operator of the mode change, as well as providing visually different information boxes depending on the current mode.

METHOD

Materials

Secondary Tasks

There were three secondary tasks used in this experiment, which a participant was expected to perform throughout the scenarios (they were not automated). The tasks occurred randomly every five to fifteen seconds, at the same rate for all secondary tasks. The three tasks were selected to help differentiate between the two theories of attention discussed in the introduction. These tasks were slight modifications of the tasks used in Schumacher, Seymour, Glass, Fenscik, Lauber, Kieras, & Meyer (2001). Two of the secondary tasks were developed to determine if MRT was supported. In this experiment, the primary task was a spatial, visual task that taxed central processing. The first secondary task was designed to stress similar resources. It was a spatial, visual task that contained a flashing arrow on the center of the display, which required the participant to choose the corresponding arrow button on the screen. The second task was designed to stress different resources. It was an auditory task, which required the participant to press either a down arrow ("v") or an up arrow ("^") button on the screen within five seconds of when a tone was heard. The "v" was used for the low tone (880 Hz) and the "^" was used for the high tone (3250 Hz). The two tasks were designed to have events occur at the same intervals during the primary task.

If the secondary task performance was worse in terms of response time and percent correct when participants were performing the similar task than when they performed the dissimilar task, MRT would appear to be supported. If the predicted effect did not occur, Neumann's view might be supported (depending on the results of the third task) or the secondary task may not be taxing enough to cause differences (although this was somewhat unlikely, given these tasks were used previously).

The third task elucidated the feasibility of Neumann's view of attention. This task was the same as the flashing arrow task mentioned above, however in this case the arrow was located in the bottom right corner of the display in the periphery (as such it would result in worse performance than the dissimilar task if MRT was correct because it taxed the same resources as the primary task). This task was used because Neumann's view stated that information could be attended to simultaneously if it was located close to the major information. There is no effect of information load. Therefore, this theory would predict that an arrow in the foveal area would always result in better performance than the arrow in the periphery, regardless of the number of targets per minute. It was not clear how the tone task should be affected, but it seems as if it would fall between the two flashing arrow tasks. In other words, the tone task was not predicted to be affected by number of targets per minute and the performance was expected to be worse than the foveal light, but better than the periphery light.

Switching of Allocations

This study required participants to interact with unmanned vehicles in a varying environment, as well as perform one of the secondary tasks. Participants were required to deal with pop-up targets. As the number of targets per minute increased, the participant's performance was anticipated to decline on the secondary task, requiring a dynamic reallocation of tasks between the participant and the vehicle (when in the dynamic mode). The specific tasks that were reallocated are found in Appendix A.

This experiment considered three levels of autonomy. They coincided with levels 5, 6, and 7 of Sheridan-Verplank's Scale, which are interesting because they represent the levels of collaboration between the human and the machine. Level 5 was the "permission" mode, in which the computer proposed an option and performed it if the human signaled approval. Level 6 corresponded to the "veto" mode, in which the computer selected an action and performed it unless the human intervened to cancel it. Lastly, the autonomous mode (level 7) occurred when the computer chose an action, performed it, and informed the human after doing so. The dynamic function allocation condition varied between all three levels of autonomy, whereas the static condition only operated at the "veto" level (Level 6), which was the intermediate level of autonomy.

Switching between levels of autonomy was an important consideration for the dynamic mode of operation. Prior research (i.e., Parasuraman, Mouloua, & Molloy, 1996) switched between the levels of autonomy when performance fell below 55% on the secondary task. While this was successful for them because they only performed one switch, this experiment was designed to have multiple levels of autonomy and switches.

Therefore, a relative performance measure like theirs could quickly result in multiple switches. For instance, if the performance fell below 55% on the secondary task and the next lowest level was switched to, the next response was critical. If the participant got that response wrong, the system would switch to the next lowest level of autonomy – quickly falling through the levels of autonomy. Clearly, that was not an appropriate system behavior. To combat these issues, this experiment used a "fixed" level of switching. In this scheme, if the participant's performance on the secondary task was between 100 and 80%, they were in the "permission" mode; between 79 and 60% performance, the level of autonomy was the "veto" mode; and below 59% performance, the level of autonomy was "autonomous". In order to determine the percentage of performance, the last ten responses to the secondary task were considered.

Another important consideration for the dynamic mode was the behavior immediately after the switch in the level of autonomy was made. Many researchers (e.g., Moray, Inagaki, & Itoh, 2000; Parasuraman, Mouloua, & Molloy, 1996) executed an immediate switch in modes and ignored any lag in the performance. However, in this experiment, there were, in most cases, still targets in the old mode visible on the screen when the targets in the new mode began appearing. Although it is clear which target belonged in which mode, there was a brief transition period with targets in both modes, before the new mode took over. This lag in the mode was normally quite brief (~5-10 seconds). Scallen, Hancock, & Duley (1995) suggested that this transition time period often exhibited different behavior than the norm. They recommended examining this "window" separately from the other mode data, as it often resulted in changes in both the primary and secondary task, with performance (in terms of response time and percent correct) on one usually being sacrificed for the benefit of the other. In this experiment, the response time and percent correct on the primary and secondary task during the window was examined, which included the old mode targets, as well as any new targets that appeared prior to the last of the old mode targets being acted upon. If, on average, the window behavior were different from the rest of the data, the results would be summarized separately. If this window showed no behavior differences from the other data, these data points would be combined with the rest of the data.

It was also important to consider the control and display requirements for the engagement and disengagement of the automation when in a dynamic allocation mode, so participants would be able to maintain situation awareness and mode awareness. Systems need to support participant monitoring by highlighting changes in goals and activities and notifying the participant of actions taken. This points to a need for transparency of the human-machine system (Olson & Sarter, 2001). In order to support this transparency, a notification indicated whenever there was a change in the level of autonomy. Also, the different levels of autonomy looked visually different, as the permission mode had a proposal with an execute button, the veto mode had a proposal, and the autonomous mode had a notification only. In addition, when in the autonomous mode, the same information was displayed; it just was not selectable or editable.

Expert System

In order to support the implementation of dynamic function allocation at multiple levels of autonomy, this study employed a knowledge-based decision aiding system.

This system should facilitate the mapping of the user's mental model, thus helping the operator to develop and maintain awareness of the system. It was implemented with an associate system, as supported by commercial software, PreAct. An associate system is a decision-aiding system that uses knowledge about the system hosting it, the environment and context in which it is operating, and the potential purposes and tasks of the system and its operators. It is a real-time decision aid that can act interactively and cooperatively with one or more humans to perform complex tasks.

The fundamental premise of this system is that, rather than viewing vehicles as being under the control of an operator, the operator and the vehicles should be viewed as collaborating agents attempting to accomplish a set of common goals. An operator commands an agent or group of agents by passing them specific high-level goals or high-level plans. Once an agent receives a plan or goal, a planning cycle is triggered that decomposes into lower level plans, goals, and actions. Frequently during the decomposition, the agent may be required to choose between alternative plans to satisfy a goal. Once the plans are decomposed into actions, these actions are executed to accomplish the objective.

Although the technology theoretically supported a dynamic shifting of responsibility (cooperative interactions), this was not empirically tested prior to this study. The dynamic switching was done by monitoring the participant's task performance and allowing the associate system to assume increasing amounts of work, as warranted by participant performance. This was done by switching between three levels of autonomy, using a performance-based change in allocation. An associate system was a good way to model this allocation method, as the system was able to monitor operator performance and the state of the system on a continual basis, which was one of the requirements for this type of system. In addition, the associate system was able to predict the participant's intent, making it a good candidate for supporting dynamic performance-based function allocation.

Simulator

The simulation was built on a Linux platform, using an existing simulator program (the Tactical Entity Simulator (TES)). This study utilized an environment designed for an earlier experiment. The mission, however, was shortened and enhanced to provide more opportunities for workload changes. In addition, in the dynamic condition, the three levels of autonomy were modified based on workload, as opposed to the original design where a participant experienced one level of autonomy throughout a trial. The participant was an operator of four Unmanned Combat Air Vehicles (UCAVs) in a simulated surveillance mission, which required them to monitor the UCAVs progress and approve and/or respond to actions based on targets that appeared. This experimental task was not studied previously, and thus further elucidated the tasks that are sensitive to changes in workload.

Simulated Domain

UCAVs are envisioned as general purpose, reusable, tactical air vehicles that can provide much lower cost and risk per delivered munitions against many classes of tactical targets than other delivery systems can. Being uninhabited, UCAVs can achieve much lower development, acquisition, and operational costs than man-rated aircraft.

A major challenge to the successful use of UCAVs is the necessity to achieve combat effectiveness and safety under the wide range of conditions that can occur on the modern battlefield. Unlike present reconnaissance unmanned vehicle operations, it will not be possible to pre-program the entire mission, nor will it be possible to continuously pilot the UCAV from a remote cockpit. Thus, a technology that provides intelligence on the vehicle to allow for autonomous behaviors, while assuring the participant has the capability to control the vehicle is highly desirable.

The use of multiple unmanned vehicles that must be monitored and directed by a participant provided a realistic task environment, while allowing the ability to realistically change the number of targets per minute to force the changing of allocations.

Subjective Mental Workload

A paper and pencil subjective workload measure was given after each of the conditions. Specifically, the NASA Task Load Index (TLX) workload measure was used to measure the participant's assessment of their workload. The NASA TLX has six subscales: mental demand, physical demand, temporal demand, performance, effort, and frustration level. The participant assessed their level on each of the scales. The TLX is designed to be given immediately after task performance because taking the TLX in the middle of a complex task can add to the workload experienced, thereby giving a contaminated workload measure (Tsang & Wilson, 1997). Research has shown that 15-30 minutes delay does not interfere with the workload ratings (Tsang & Wilson).

Three of the scales in the TLX reflect demands on participant resources (mental, temporal, and physical demands), and three characterize interaction between the participant and the task (effort, performance, and frustration) (Warm & Dember, 1998). Of these, mental workload and frustration appear to be the primary determinants of workload (Warm & Dember).

Situation Awareness

A paper and pencil self-rating situation awareness measure was also given after each of the conditions. This is a subjective measurement, which is advantageous because it is easy and inexpensive to administer, and allows participants to fill it out after task performance. This is unlike the objective measures (e.g., SAGAT), which require the scenario to be stopped mid-task, possibly affecting task performance and workload measurements. Specifically, the Situation Awareness Rating Technique (SART) (Taylor, 1990) was used to measure the participant's assessment of their situation awareness. It measures the amount of demand on attentional resources and the understanding of the situation provided. This tool was developed using aircrew constructs of situation awareness, therefore should be an ecologically valid measure for the proposed task (Selcon & Taylor, 1990; Taylor, 1990).

A full description of the development of the SART, including elicitation of the constructs, rating of the constructs, and a principal components analysis of the constructs, is available in Taylor (1990). This analysis revealed there are ten basic dimensions of the SART, which can then be grouped into three components: demands on attentional resources, supply of attentional resources, and understanding of the situation. Thus, the administered SART has fourteen dimensions (ten dimensions, the three aggregate components, and an overall SA measure), which the participant is asked to rate their awareness on a scale of 1-10. After these self-ratings are collected, they are combined into a measure of participant SA with the equation (Taylor, Shadrake, Haugh, & Bunting, 1996):

$$SA = \sum U/n_u - (\sum D/n_d - \sum S/n_s)$$

This equation has the measures of understanding added together divided by the number of measures. The difference of the demand and supply are then subtracted from this amount. This equation shows that understanding decreases when the demand is greater than the supply, but increases when supply is greater than demand.

Taylor (1990) described the validation of the SART using a different group of subjects then were originally used in the construction. This resulted in statistical significance of all dimensions except information quality. Similarly, Selcon and Taylor (1990) assessed the SART in three experiments and found the same three major factors. In addition, the factors appeared to be sensitive to manipulations of independent variables, with demand increasing with increased number of tasks, understanding increasing with a more rule-based task, and supply increasing when using a continuous task. Increased understanding also was found to reduce response times. Later work by

Vidulich, McCoy, and Crabtree (1996) found that increasing the complexity of a flying task resulted in increased demand, whereas increasing the SA information available to the participant resulted in increased understanding. Thus, the sensitivity and diagnosticity of the SART was demonstrated in a variety of different situations.

Participants

This study utilized sixty participants, ages 18-28 years. There were three experimental groups, described in the secondary task section, with 20 people assigned to each group. There were no significant differences between the groups in terms of demographics; which are described in Table 2 below. The questions were standard questions for demographics questionnaires, with the exception of the video game question. This was added because this task is somewhat similar to a video game, so differences between the groups on this question could have affected the results. The participants were students attending the Georgia Institute of Technology, enrolled in psychology courses. All of the participants received either course credit or $15 for their participation. All participated in one 1.5-hour session, including a 30-minute training session and two 25-minute experimental tasks. The remainder of the time was used for filling out paperwork: before, during, and after the experiment.

Table 1. Demographic Data (means and standard deviations)

	Foveal		Periphery		Tone	
Age	21.32	(3.23)	21.75	(3.74)	22.50	(3.22)
Gender [a]	1.85	(0.37)	1.75	(0.44)	1.55	(0.51)
Education (years) [b]	3.70	(1.13)	3.75	(1.25)	3.65	(1.14)
Video Games [c]	4.90	(1.59)	5.55	(1.15)	5.20	(1.70)
Marital Status [d]	1.10	(0.31)	1.10	(0.31)	1.15	(0.49)
Race [e]	2.65	(0.59)	2.85	(0.49)	2.70	(1.53)

a. A gender of 1 indicates females and a gender of 2 indicates a male
b. Score of 3 indicates some college experience
c. Frequency of playing video games is between 1 (4 hours a day) and 7 (never)
d. A score of 1 indicates never married and 2 indicates married
e. 1 indicates Asian, 2 indicates African American, 5 indicates Caucasian

Design

This study utilized a 2 (adapting type) X 3 (number of targets per minute) X 3 (secondary task) design, with adapting being a within subject variable, number of targets per minute being a within-subject variable, and secondary task being a between subject variable (so there was no interference between the tasks). There were twenty people in each secondary task, for a total of 60 people. The three levels of secondary task were discussed in the previous section. The adapting type was whether the participant experienced static or dynamic function allocation during that condition. The conditions were equated on the number of targets and length of the scenario, and the order of presentation of the static or dynamic mode was randomized to prevent order effects.

All participants experienced three levels of the number of targets per minute while the vehicles were in the mission area. This was done by breaking the scenario into six 2-minute time blocks, which varied in number of targets per minute (see figure below for a sample). The order of presentation of the low, medium, and high targets per minute was randomized for the participants. The low number of targets per minute condition had one event per minute, the medium condition had 3 events per minute, and the high condition had 5 events per minute. This resulted in a total of 36 events per scenario.

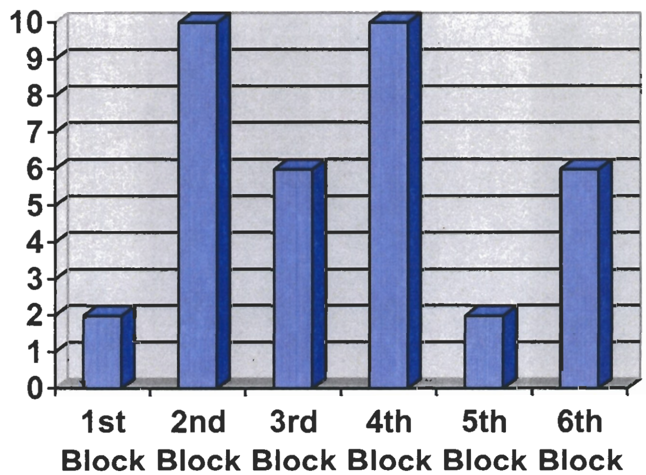

Figure 10. Time blocks of the number of targets per minute

There were four dependent measures: response time, percent correct, situation awareness, and subjective mental workload. Both response time and percent correct were measured for the primary and secondary tasks. For the primary task, response time and percent correct were computed using responses from when the operator was performing in the permission and veto modes (the autonomous mode was considered missing data for the data analyses), as these represented the operator's performance on this task. Response time was measured from the time a target had a proposal until the participant dealt with it. Response time was not a completely equated measure when one considers that the permission mode had two actions and the veto mode had only one action. However, given that both attention theories predict that the dynamic condition will be slower at the lower LOA, this did not bias against either theory. In addition, the permission mode was supposed to involve higher workload and/or more actions, so this seemed to be an appropriate way to distinguish between the modes.

Percent correct was assessed by determining whether the correct action was taken. In this task, any target inside the Area of Operation (outlined by a box) could be shot at, whereas any target outside the area should not be attacked. The decision aid proposed an attack for all targets, so the participant was supposed to reject inappropriate attacks. Half of the targets in each block were outside the Area of Operation, for a total of 18 targets. While this made the automation unreliable according to some researchers, it did this for both modes and for the training. So, even if the participants had problems with the unreliability (i.e., slower response times), it occurred for all modes and training. This would not bias against either mode or attention theory. In addition, this condition was termed imperfect/aware, in that participants were aware that the automation was not perfectly reliable (Wickens, 2000). When operating in this environment, participants have been shown to develop strategies in which they allocate a portion of their resources to considering alternative hypotheses and searching for cues that might support those hypotheses (Wickens). Thus, the unreliable automation makes participants verify the suggestions from the automation before making responses, which should actually prevent automation biases and ensure more accurate results (although probably slower responses).

For the secondary task, the operator performed the task during all three modes, so response time and percent correct were measured throughout task performance. Response time was the time from when the arrow appeared or the tone sounded until the participant chose a button. Percent correct was determined by whether the correct button was selected. If the participant did not respond to the secondary task, the response was considered incorrect. Situation awareness was assessed with the SART tool, as described above. Workload was assessed with the NASA TLX, as described above.

Procedures

Upon starting the experiment, the participant was given a packet of paperwork containing two copies of the experimental consent form and the demographics form. When they completed the paperwork, the participants were given a 30-minute training session to familiarize themselves with the task and the displays. Although the training period was probably too short to develop any differing strategies for interacting with the static and dynamic FA conditions, the training was done as a mixed condition training session (both dynamic and static FA were demonstrated within a single training session). This should have prevented participants from developing different strategies, so the results were more likely to represent experimental effects rather than strategy differences (see Strayer & Kramer, 1994 for a discussion of mixed versus blocked training). Participants were instructed to weight the primary task more than the secondary task. After training, participants assessed their workload using the NASA TLX and their situation awareness using the SART, so they could familiarize themselves with the measures.

The training was followed by the first condition of the experiment (either static or dynamic function allocation). The order of presentation was counterbalanced to prevent order effects, such that half the participants were randomly assigned to the static allocation scenario first and the rest did the dynamic allocation scenario first. The participant then assessed their situation awareness and workload. After a short break, they did the same sequence with their second condition. This ordering was designed to minimize the interference between the conditions of the task and allow the experimenter to get an assessment of situation awareness and workload for each level of adapting.

Each run through the scenario began when the participant indicated they were ready to begin. The simulation then initialized, vehicles began appearing, and the secondary task started. As the scenario continued, targets begin appearing at differing rates. At the conclusion of the condition, the participant filled in the SART and NASA TLX for that condition. The simulation was shut down and reset so the next condition was ready when the participant indicated they were ready.

At the conclusion of both of the scenarios, the participant filled out an exit interview and received a debriefing, which explained the purpose of the experiment and elicited any feedback from them.

Predicted Results

This discussion deals with the dependent variables: percent correct, response time, situation awareness, and workload. A significant difference was expected between the mean secondary task percent correct scores of the participants in the dynamic versus static function allocation methods. This was expected because the participants in the dynamic mode were predicted to be actively involved with the system at lower levels of autonomy and have less workload at higher levels of autonomy, theoretically producing better performance. For the primary task, the dynamic mode was expected to result in worse performance than the static mode until the highest number of targets per minute because there were more actions to carry out. There were no specific predictions between the different secondary tasks or for interactions involving the secondary tasks. Increased number of targets per minute was not expected to decrease the percent correct of the secondary task in the dynamic allocation, as it should compensate for participant deficiencies, by switching levels of autonomy. In the static allocation, percent correct of the secondary task was expected to decrease with increased number of targets per minute. Thus, a number of targets per minute by adapting type interaction was expected for percent correct. There were no specific predictions for an interaction between adapting type and number of targets per minute for primary task percent correct. Similar results were expected for response time.

For the dependent variable situation awareness, there were no specific effects are predicted for the secondary tasks. The adapting type was predicted to show a significant effect for situation awareness, as the dynamic task allocation was predicted to result in better situation awareness than the static allocations. Similar results were expected for the workload assessment although the direction of the interaction would be opposite (high SA is related to low workload). For a full explanation of the predicted effects, see the table below.

Table 2. Predicted effects

	Response Time	Percent Correct	WL	SA
Adapting	Dynamic faster for secondary task- more attention available. For primary task, dynamic slower than static until highest number of targets	Dynamic more accurate for secondary task- more attention available. For primary task, dynamic less accurate than static until highest number of targets.	Dynamic lower - workload should be more constant with the changing allocations	Dynamic higher - more attention available to maintain SA
Number of targets per minute	More targets - slower performance	More targets - less accurate performance	More targets - higher workload	More targets - lower SA
Secondary Task	If MRT, then ordering is dissimilar, then periphery & foveal & if Neumann, foveal, periphery, dissimilar	If MRT, then ordering is dissimilar, then periphery & foveal & if Neumann, foveal, periphery, dissimilar	If MRT, then ordering is dissimilar, then periphery & foveal & if Neumann, foveal, periphery, dissimilar	If MRT, then ordering is dissimilar, then periphery & foveal & if Neumann, foveal, periphery, dissimilar
A X N	Task more diff. -- dynamic the same, static gets slower (MRT) OR static the same and dynamic getting faster (Neumann) for primary task, secondary task - dynamic the same, static getting slower	Task more diff. -- dynamic the same, static gets less accurate (MRT) OR static the same and dynamic gets more accurate (Neumann) for primary task, secondary task - dynamic the same, static gets less accurate	N/A	N/A
A X S	No effect predicted - would indicate different reaction of secondary task (attention theory) when in different adapting mode	No effect predicted - would indicate different reaction of secondary task (attention theory) when in different adapting mode	No effect predicted - would indicate different reaction of secondary task (attention theory) when in different adapting mode	No effect predicted - would indicate different reaction of secondary task (attention theory) when in different adapting mode
S X N	If MRT, interaction predicted, where tone less affected than foveal. If Neumann- no interaction predicted	If MRT, interaction predicted, where tone less affected than foveal. If Neumann- no interaction predicted	N/A	N/A
A X S X N	No effect predicted - prediction would show difference in reactions of the secondary tasks (attention theories) when increasing difficulty in the static versus dynamic conditions	No effect predicted - prediction would show difference in reactions of the secondary tasks (attention theories) when increasing difficulty in the static versus dynamic conditions	N/A	N/A

RESULTS

The statistical analyses of the data consisted of a Multivariate Analysis of Variance (MANOVA) and a three-factor analysis of variance (ANOVA). Both of these analyses were mixed model types, with adapting type and number of targets per minute as within subject variables and secondary task as a between subject variable. Significant interactions were examined using post hoc tests of partial interaction contrasts. In addition, there were planned comparisons to examine the attention theory predictions. An alpha level of .05 was assumed for all statistical tests.

Window Behavior

The first important analysis was whether the window behavior was different from the rest of the data. The window was defined as the time after a mode transition until all of the old mode decisions had been cleared from the screen. If the results were different, these window data would be summarized separately. If not, these data points would be combined with the rest of the data. When examining the windows for primary and secondary task response time and percent correct, there were no significant differences between the windowed behavior and the rest of the data, thus these data points were combined with the rest of the data. The rest of the analyses were done with all of the data.

A summary of the means and standard deviations for conditions and dependent variables is shown in Table 3.

Table 3. Summary of Dependent Variables (means and standard deviations)

	Static Mode	Dynamic
Primary Task Response Time – Low (sec)	.33 (.33)	.59 (.64)
Primary Task Response Time – Medium (sec)	.62 (.64)	1.38 (1.08)
Primary Task Response Time – High (sec)	.58 (.49)	1.12 (.89)
Primary Task Percent Correct – Low (%)	87 (17)	87 (18)
Primary Task Percent Correct – Medium (%)	83 (16)	80 (13)
Primary Task Percent Correct – High (%)	77 (16)	82 (12)
Secondary Task Response Time – Low (sec)	.07 (.09)	.08 (.11)
Secondary Task Response Time – Medium (sec)	.08 (.10)	.08 (.09)
Secondary Task Response Time – High (sec)	.10 (.10)	.11 (.12)
Secondary Task Percent Correct – Low (%)	86 (20)	79 (24)
Secondary Task Percent Correct – Medium (%)	70 (19)	66 (25)
Secondary Task Percent Correct – High (%)	57 (28)	61 (98)
Total WL (out of 100)	56.29 (15.98)	58.28 (14.00)
Mental Demand (out of 100)	68.42 (20.86)	70.13 (18.80)
Physical Demand (out of 100)	28.66 (23.85)	29.83 (23.37)
Temporal Demand (out of 100)	70.38 (23.33)	69.88 (20.26)
Performance (out of 100)	47.67 (23.47)	53.38 (20.18)
Effort (out of 100)	66.33 (20.03)	69.04 (16.73)
Frustration (out of 100)	56.25 (22.99)	57.46 (24.79)
Total Situation Awareness (out of 100)	52.87 (22.92)	51.54 (22.29)
Cognitive Demand (out of 100)	65.83 (20.52)	66.29 (21.05)
Instability of Situations (out of 100)	63.54 (20.90)	64.79 (21.88)
Complexity of Situations (out of 100)	48.58 (23.86)	49.13 (25.49)
Variability of Situations (out of 100)	62.75 (21.46)	67.46 (21.49)
Supply of Resources (out of 100)	58.38 (22.00)	52.96 (20.99)
Readiness (out of 100)	51.83 (23.13)	53.75 (24.18)
Concentration of Attention (out of 100)	55.88 (22.19)	56.08 (24.35)
Division of Attention (out of 100)	54.79 (19.81)	51.79 (23.29)
Spare Mental Capacity (out of 100)	41.33 (23.41)	38.71 (22.14)
Understanding of Situation (out of 100)	61.42 (23.54)	62.50 (23.28)
Quantity of Information (out of 100)	54.95 (21.47)	55.04 (23.31)
Quality of Information (out of 100)	57.88 (23.40)	56.96 (22.29)
Familiarity (out of 100)	56.75 (25.01)	58.25 (26.54)

Primary task response time

When examining the primary task performance data using a MANOVA, there was no speed-accuracy tradeoff for primary task response time. After examining the MANOVA, an ANOVA was done for primary task response time. These analyses revealed a significant main effect for adapting type, $F(1,57) = 29.000$, $p < .001$, $MSE = 24.531$, where the participants performed significantly faster in the static condition than in the dynamic condition. There was also a significant main effect of the number of targets per minute, $F(2,56) = 55.368$, $p < .001$, $MSE = 9.273$, in which the response time when participants experienced the low number of targets per minute was significantly faster than the response time at either of the other two levels of targets per minute. Both of these findings were as predicted.

In addition to the main effects, there were two significant interactions. The first was a secondary task by adapting type interaction, $F(2,57) = 3.748$, $p = .030$, $MSE = 3.170$. This interaction indicated that, for response time, performance in the static mode was not significantly affected by the secondary task, whereas for performance in the dynamic mode, the tone task resulted in significantly slower response times than the foveal task, $t(38) = 2.302$, $p = .027$. The other secondary task comparisons were not significant (see Figure 11). This finding was contrary to MRT, in that the foveal task was predicted to result in slower response times than the tone task. Neumann's theory did not directly address the tone task, however the prediction was that the foveal task would result in faster performance than the tone task, but the performance on the periphery task would be worse than both. Thus, neither theory was completely supported by this interaction.

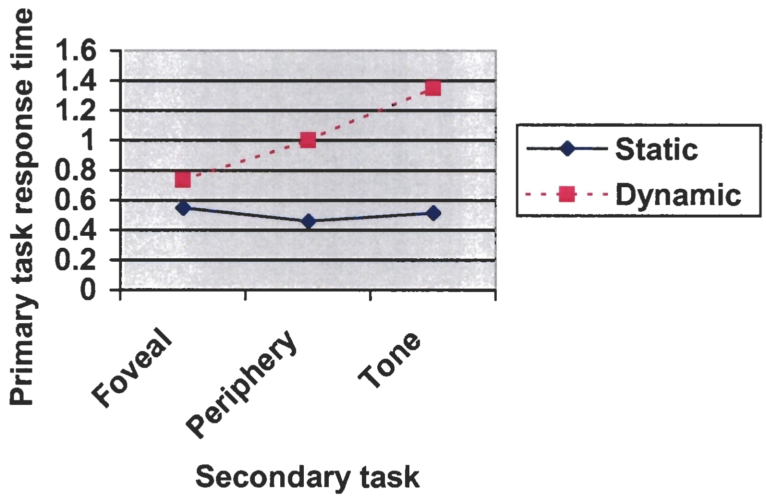

Figure 11. Interaction between adapting type and secondary task for primary task response time

The second significant interaction was between the adapting type and the number of targets per minute, $F(2,56) = 18.502$, $p < .001$, $MSE = 1.866$. This interaction is discussed more in a later section that discusses attention theories and their predictions (see Figure 12).

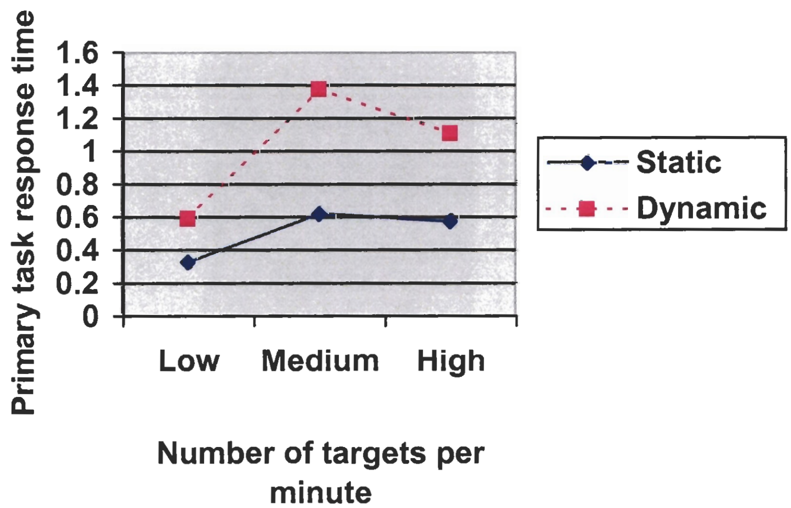

Figure 12. Interaction between number of targets per minute and adapting type for primary task response time

Thus, for response time there was no interaction between the number of targets per minute and the secondary tasks. This is supportive of Neumann's theory of attention, whereas MRT predicted that the tone task would be less affected than the foveal task by increasing the number of targets per minute. For primary task response time, MRT was not supported, whereas Neumann's theory was supported by the two interactions.

Primary task percent correct

Similar to primary task response time, no speed-accuracy tradeoff for primary task percent correct was revealed with a MANOVA. An analysis of the percent correct for the primary task also revealed a number of significant effects. The first, a main effect of the number of targets per minute, $F(2,56) = 11.121$, $p < .001$, $MSE = .236$, showed that performance was significantly more accurate at the low number of targets per minute than at the medium, $t(59) = 2.621$, $p = .011$, or high number of targets per minute, $t(59) = 3.973$, $p < .001$. In addition, the performance during the medium number of targets was significantly more accurate than performance at the high number of targets per minute, $t(59) = 2.106$, $p = .039$. This main effect was predicted.

There were also two interaction effects. The first was an interaction between the type of secondary task and the number of targets per minute, $F(4,114) = 4.573$, $p = .002$, $MSE = .097$. This interaction showed that for percent correct performance on the foveal task, there was not a significant change at any number of targets per minute, whereas the performance in the periphery and the tone tasks declined significantly as the number of targets per minute increased. For the primary task percent correct when a participant experienced the periphery task, there was a significant difference between the low and medium number of targets per minute, $t(19) = 2.541$, $p = .021$, and the low and high number of targets per minute, $t(19) = 3.737$, $p = .001$. The participants who experienced the tone task showed significant differences between all three levels of number of targets per minute (see Figure 13). The existence of this interaction was counter to Neumann's predictions. However, it was also counter to MRT, in that the foveal task was the least affected by increasing the number of targets per minute. At the low number of targets, MRT appeared to be supported, with the ordering of secondary tasks being tone, periphery, foveal, however this reversed as the number of targets increased.

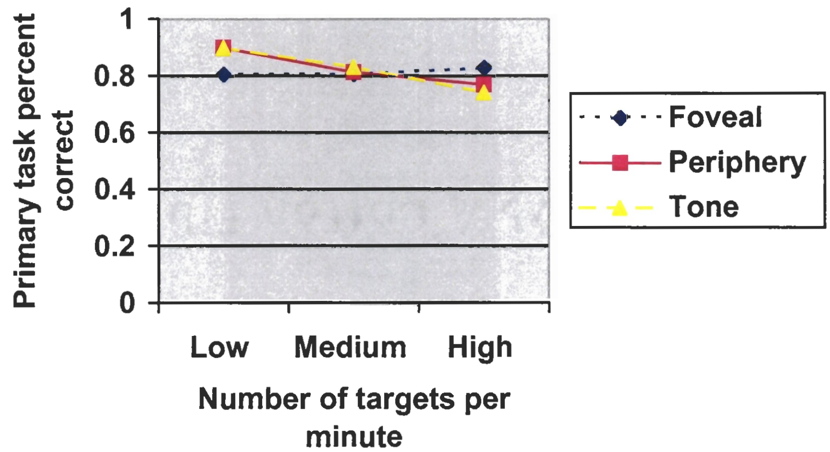

Figure 13. Interaction between the number of targets per minute and secondary task for primary task percent correct

The second interaction shown by the analysis of the percent correct for the primary task was between the adapting type and the number of targets per minute F(2,56) = 4.493, p = .013, MSE = .111. This interaction is also discussed later in this section (see Figure 14).

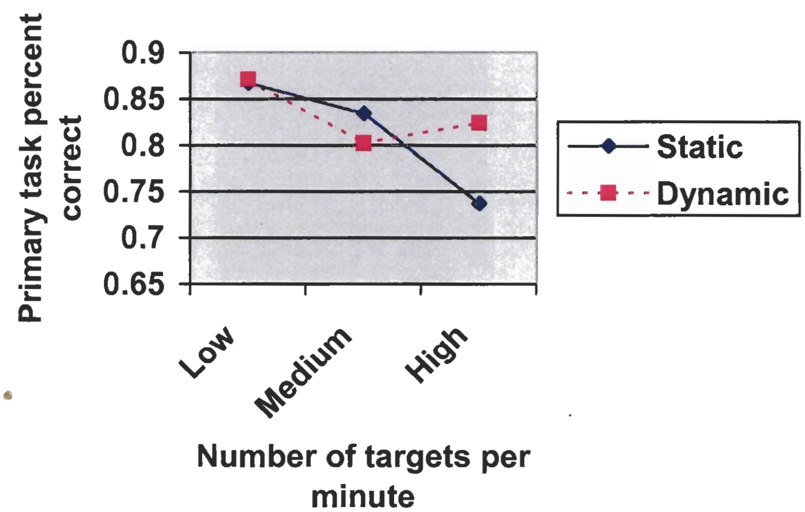

Figure 14. Interaction between number of targets per minute and adapting type for primary task percent correct

Secondary task response time

In general, the participants were able to successfully interact with the secondary tasks, even when primary task performance became demanding. Performance declined with increased demands, but participants were not getting things wrong, rather they just were not always able to respond the task, as is shown in both static and dynamic percent correct data. However, they were not simply ignoring the secondary task. This was evidenced, in the dynamic mode, in the fact that participants were able to "cycle" through the different levels of autonomy as task demands increased and decreased. If they had ignored the secondary task when experiencing the higher levels of autonomy, it would have taken them much longer for their percent correct on the secondary task to become high enough to foster a primary task mode change to a lower level of autonomy. Thus, they would have spent a large majority of their time in autonomous mode, which was not shown in the data (20% was the highest percentage of time spent in the autonomous mode).

There was a covariance effect between secondary task percent correct and secondary task response time. When examining response time and percent correct together, there was an interaction between adapting type and number of targets per minute, $F(2,38) = 5.213$, $p = .010$. This interaction revealed no speed-accuracy tradeoff, rather both the response time and percent correct performance declined as the number of targets per minute increased for the static mode, but not for the dynamic mode. This is consistent with predictions that the dynamic mode has more resources available to the participant as the number of targets per minute increase, so the secondary task performance in the dynamic mode should not decline as rapidly as the static mode. It also indicated that the participants were paying less attention to the secondary task than the primary task, which was what they were instructed to do.

For secondary task response time alone, there was a significant effect of the number of targets per minute, $F(2,56) = 38.132$, $p < .001$, $MSE = .019$, where the response time on the secondary task was fastest during the lowest number of targets per minute and slowest during the highest number of targets per minute. There was a significant difference between all three levels of targets per minute (see Figure 15), as was predicted.

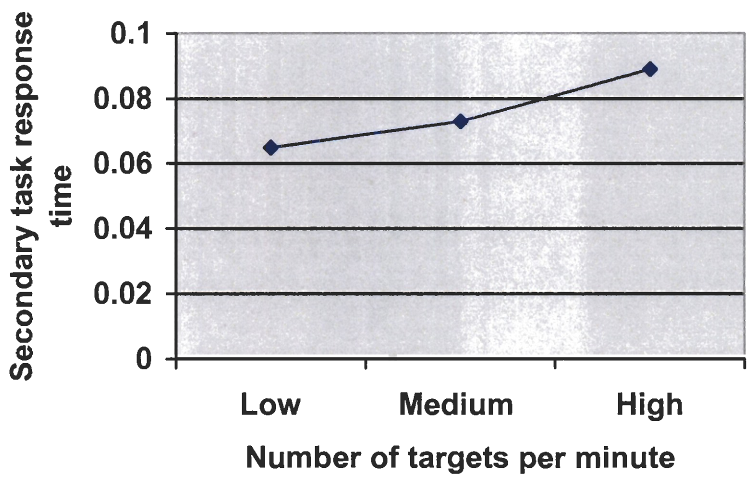

Figure 15. Number of targets per minute for secondary task response time

There was also a main effect of the secondary task, $F(2,57) = 3.152$, $p = .050$, $MSE = .042$. This effect showed that the response time performance in the periphery task was significantly slower than the response times for the tone task, $t(38) = 2.262$, $p = .030$. There was also a significant difference between the response times for the foveal and tone tasks, $t(38) = 2.836$, $p = .007$, with performance during the tone task being significantly faster than performance during the foveal task. There were no significant differences between the response times for the periphery and foveal tasks (see Figure 16). This main effect is consistent with MRT, in that performance during the tone task was significantly faster than performance in either of the other two tasks.

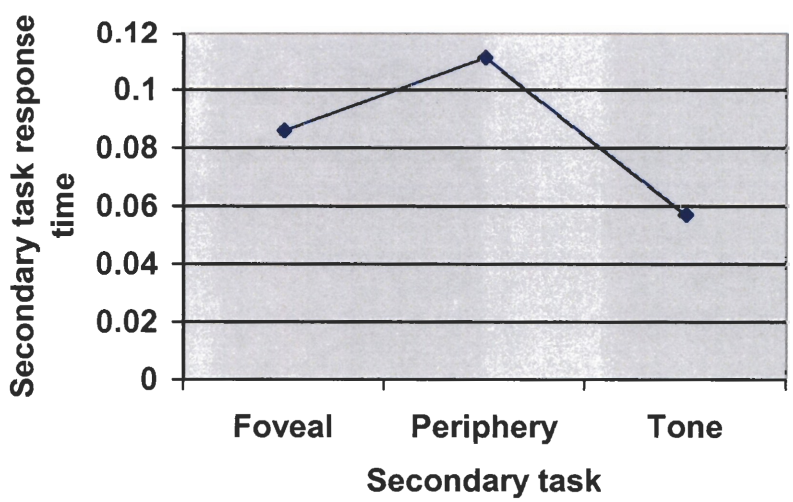

Figure 16. Secondary task response time for each level of secondary task

No significant interactions were found for secondary task response time. Because there was no interaction between secondary task and number of targets per minute, Neumann's theory would be supported.

Secondary task percent correct

For secondary task percent correct, there was a significant three-way interaction, $F(4,112) = 3.306$, $p = .013$, $MSE = 0.082$, in which the performance in the foveal task was significantly more accurate than the performance in the periphery task in the static mode at high numbers of targets per minute (although both were less accurate than the tone task). This order was reversed for the dynamic mode (see Figure 17). This was not predicted by either attention theory. However, given that the tone task was superior in both cases, this would tend to support MRT.

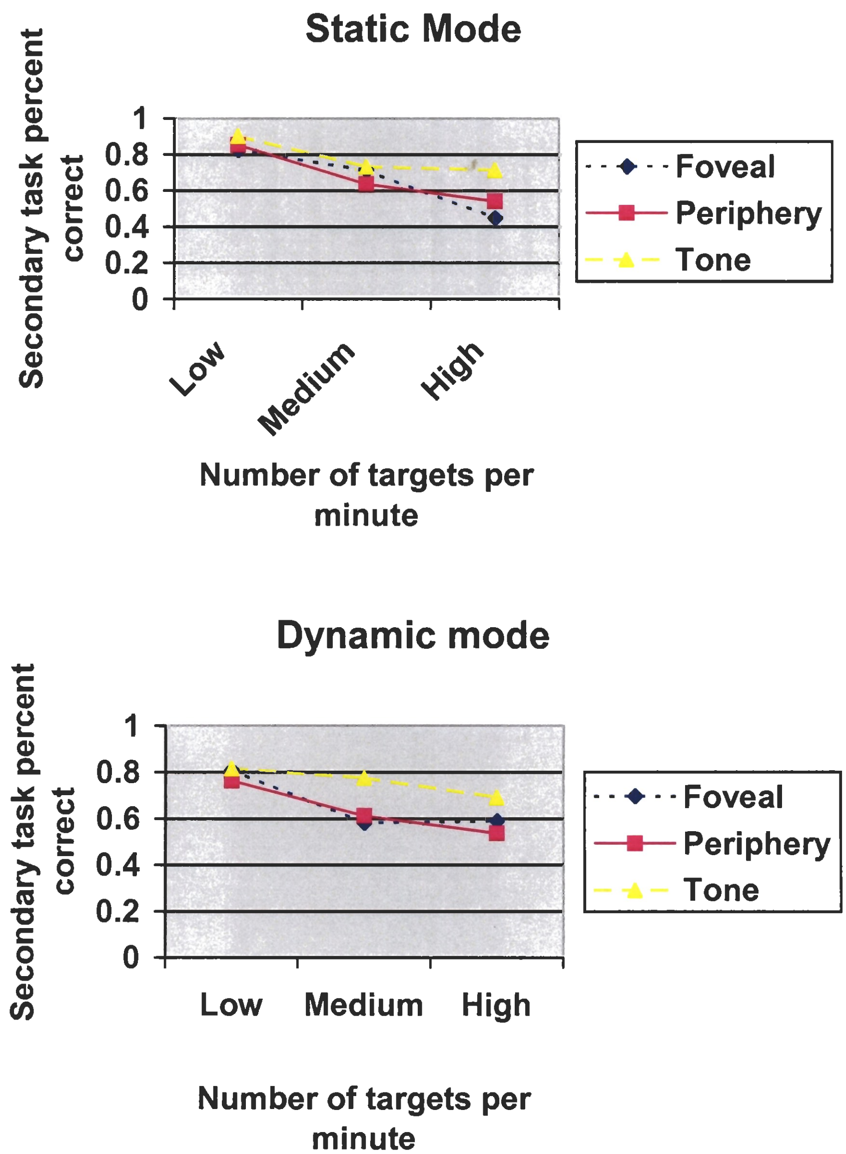

Figure 17. Three-way interaction between number of targets per minute, adapting type, and secondary task for secondary task percent correct

There was also a main effect of the number of targets per minute, $F(2, 56) = 54.719$, $p < .001$, $MSE = 1.772$, as was predicted. This effect showed that percent correct on the secondary task was significantly higher at low number of targets per minute than at either medium, $t(59) = 6.378$, $p < .001$, or high number of targets per minute, $t(59) =$

4.045, p < .001. There was also a significant difference between medium and high targets per minute, t(59) = 9.613, p < .001 (see Figure 18).

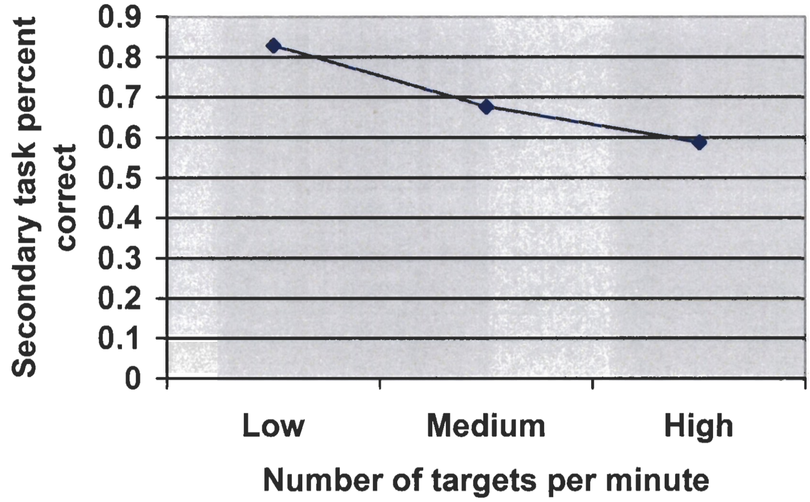

Figure 18. Secondary task percent correct for the number of targets per minute

There was another main effect, this one for the type of secondary task, $F(2, 56) = 3.442$, $p = .039$, $MSE = .515$. This main effect showed a significant difference between the tone task and the periphery, $t(38) = -2.370$, $p = .023$, where the participants interacting with the tone task had significant more accurate secondary task performance than those interacting with the periphery task. There were no differences between the other secondary tasks.

For percent correct on the secondary task, there was no significant interaction between number of targets per minute and the secondary tasks. Again, this would lend support to Neumann's predictions.

Workload and Situation Awareness

When examining the workload and situation awareness measures, the results were not as telling. For workload, there were no significant main effects or interactions for total workload, mental demand, physical demand, temporal demand, performance, frustration, and effort. This suggests that either workload did not change significantly throughout task performance or the NASA TLX was not an effective measure of workload for this task. However, given that performance did change throughout the task, it seems more likely that the NASA TLX was not an effective tool. When examining the situation awareness measures, there were no significant main effects or interactions for total situation awareness, demand on cognitive resources, complexity of situations, readiness, concentration of attention, division of attention, spare mental capacity, understanding of the situation, information quantity, information quality, and familiarity of the situation.

There was a significant interaction between adapting type and secondary task for instability of situations, $F(2, 57) = 3.914$, $p = .026$, $MSE = 1254.531$. This showed no significant differences between the three secondary tasks in the ratings of instability of the situation for the dynamic mode. However, in the static FA, the tone task was rated as significantly more instable than the periphery task (see Figure 19). This finding was somewhat unexpected, considering there are no mode changes in the static condition, so it is strange that there should be differences in the instability ratings for the static FA. However, it does support the predictions that the performance in the dynamic condition should have less instability than the static condition. On the other hand, the participants rated the static mode as less instable than the dynamic mode for the periphery task, although not significantly. This was counter to predictions, as the dynamic mode was predicted to represent higher levels of situation awareness (and less instability) for all secondary tasks.

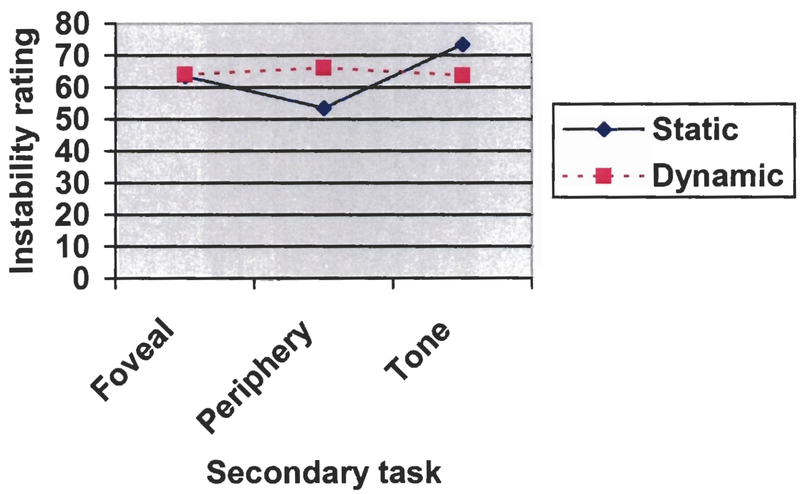

Figure 19. Interaction between secondary task and adapting type for instability of situations

There was also an adapting type by secondary task interaction for variability of situations, F(2,57) = 3.460, p = .038, MSE = 943.177. In this interaction, there was a significant difference between ratings of the variability of the situation in the static and dynamic FA for the periphery task, t(19) = -2.656, p = .016, whereas the other two secondary tasks showed no differences between the FA conditions (see Figure 20). This was counter to predictions in that the performance in the dynamic condition was more variable than performance in the static condition for the periphery task. These ratings were expected to show lower variability ratings for the participants in the dynamic task, as this condition was predicted to represent higher levels of situation awareness. However, given that the dynamic adapting type had mode changes and the static one did not, it is not surprising that the participants would rate it as more variable. Perhaps this particular scale of the SART is not a good indication of SA for this task.

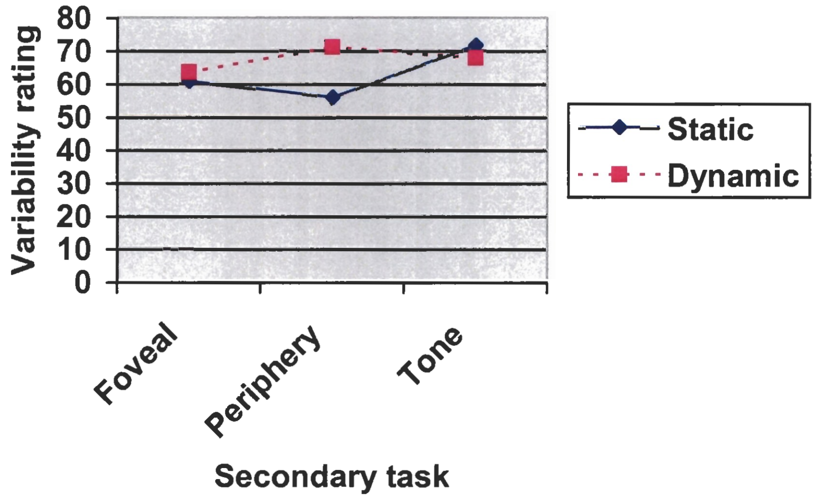

Figure 20. Interaction between adapting type and secondary task for variability of situations

One possibility for why the participants found the periphery task to be more instable and more variable in the dynamic condition was that they experienced significantly higher numbers of mode transitions in this secondary task, relative to the other two tasks. However, an examination of the number of mode transitions in the secondary tasks revealed no significant differences between them.

Another possibility was that the periphery task was the most demanding secondary task. This would be supportive of Neumann's predictions, where the periphery task was predicted to be the worst of the secondary tasks, as well as partially supportive of MRT, as the foveal and periphery task as predicted to show equally poor performance, relative to the tone task.

There was also a main effect of adapting type for supply of cognitive resources, $F(1,57) = 5.075$, $p = .028$, $MSE = 880.208$. In this case, the dynamic condition was rated as requiring significantly more cognitive resources than the static condition. This is counter to predictions, as the dynamic function allocation has been touted as requiring fewer cognitive resources at high levels of autonomy (as the system takes over some of the tasks), which provides more resources to maintaining SA. The additional cognitive resource demands experienced by participants in the dynamic condition may indicate additional effort related to determining what mode is currently active or they may be because the dynamic FA requires participants to spend more time in the permission mode, which is more demanding than the veto mode.

Analyzing the Predictions of the Research Theories

Primary task response time

When examining the two research theories and their predictions about primary and secondary task performance, there were some interesting results. For the primary task response time, there was a significant difference between static and dynamic function allocation at low, $t(59) = -2.922$, $p = .005$, medium, $t(59) = -5.712$, $p < .001$, and high numbers of targets per minute, $t(59) = -5.002$, $p < .001$. In all cases, the response time for the static condition was significantly faster than the response time for the dynamic condition, as was predicted by both theories, although the differences at the highest number of targets per minute were not predicted to be significant.

Differences within each of the function allocation modes were examined next. For the dynamic condition, performance in the low number of targets per minute was significantly faster than either performance in the medium, $t(59) = -8.545$, $p < .001$, or high, $t(59) = -6.532$, $p < .001$. Response time performance in the medium number of targets per minute was significantly slower than performance in the high number of targets per minute, $t(59) = 3.752$, $p < .001$. Similarly, within the static condition, there were significant differences in the response time for the primary task between the low number of targets and both the medium, $t(59) = -5.588$, $p <.001$, and high numbers of targets per minute, $t(59) = -5.780$, $p <.001$. In both cases, the participants interacting with the low number of targets were significantly faster than when they interacted with either of the other two levels (see Figure 21).

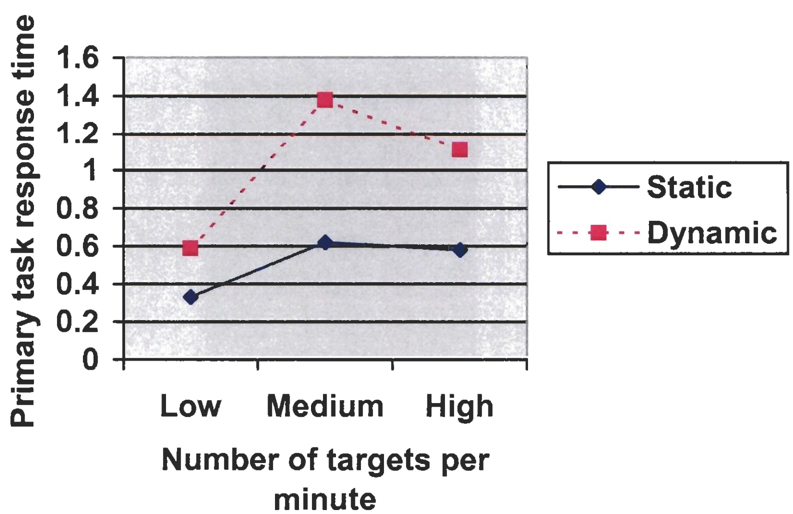

Figure 21. Adapting type by number of targets per minute for primary task response time

Interestingly, this pattern lends some support to both attention theories. MRT predicted a decline in performance for the static FA, which was supported between the low and medium targets per minute. However, from medium to high number of targets per minute, there was no significant difference in the static condition, which is supportive of Neumann's theory. In addition, from medium to high number of targets per minute, there was an improvement in the performance in the dynamic FA, which again supports Neumann's view. Thus, for primary task response time, Neumann's theory appeared to be supported at the higher levels of task demands, whereas MRT was supported at the lower levels of task demand.

Primary task percent correct

For percent correct on the primary task, there was only a significant difference between the adapting types at the high numbers of targets per minute, $t(59) = -3.258$, $p = .002$. In this case, performance in the dynamic condition was more accurate than the static condition, which was counter to predictions, as the dynamic performance was expected to approach, but not exceed static performance.

Within the dynamic condition, there was a significant difference in percent correct performance on the primary task between low and medium number of targets per minute, $t(59) = 2.464$, $p = .017$, where the low number of targets resulted in significantly more accurate performance than the medium number of targets per minute. For the static mode, there were significant differences between the high number of targets and the low, $t(59) = 4.027$, $p < .001$, and medium ones, $t(59) = 3.376$, $p = .001$. In both cases, performance in the high number of targets per minute was significantly less accurate than at the other two levels (see Figure 22).

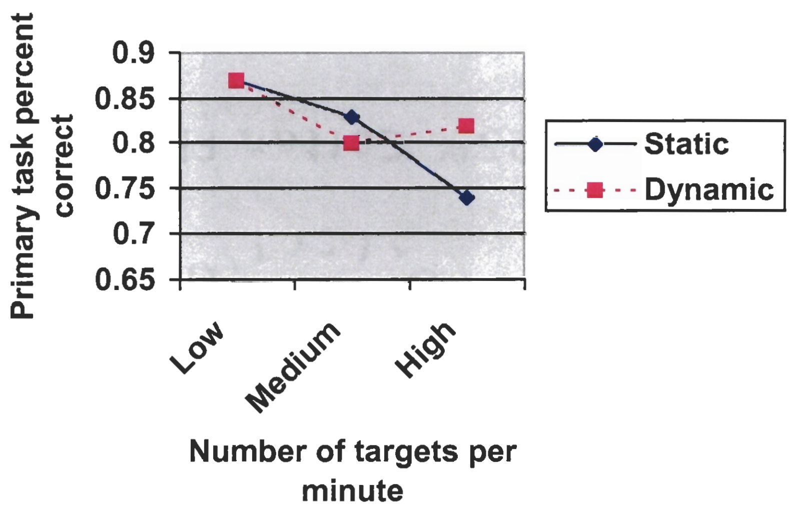

Figure 22. Adapting type by number of targets per minute for primary task percent correct

With regards to the attention theories, Neumann's theory was not supported by either the static or dynamic performance, although there was an improving trend between medium and high number of targets per minute in the dynamic mode. This trend, if significant, would have supported Neumann's theory. MRT, on the other hand, predicted a decline in the static performance, which was demonstrated. It also predicted stable performance in the dynamic mode, which was supported between the medium and high number of targets per minute.

Secondary task response time

There were no significant differences in response times on the secondary tasks between static and dynamic function allocation at any of the three levels of numbers of targets per minute. Within the dynamic condition, there was a significant difference in secondary task response time between the medium and high number of targets per minute, $t(59) = -2.767$, $p = .008$, and the low and high number of targets, $t(59) = -4.884$, $p < .001$. In both cases, the smaller numbers of targets per minute resulted in faster performance. On the other hand, the static condition showed significant differences in response time to the secondary task between each of the number of targets. For the difference between the low and medium levels, the effect showed that performance at the low numbers of targets performance was significantly faster than performance at the medium number of targets per minute, $t(59) = -4.046$, $p < .001$. Medium resulted in significantly faster secondary task performance than the high numbers of targets, $t(59) = -4.484$, $p < .001$, and low resulted in faster performance than the high, $t(59) = -5.970$, $p < .001$ (see Figure 23).

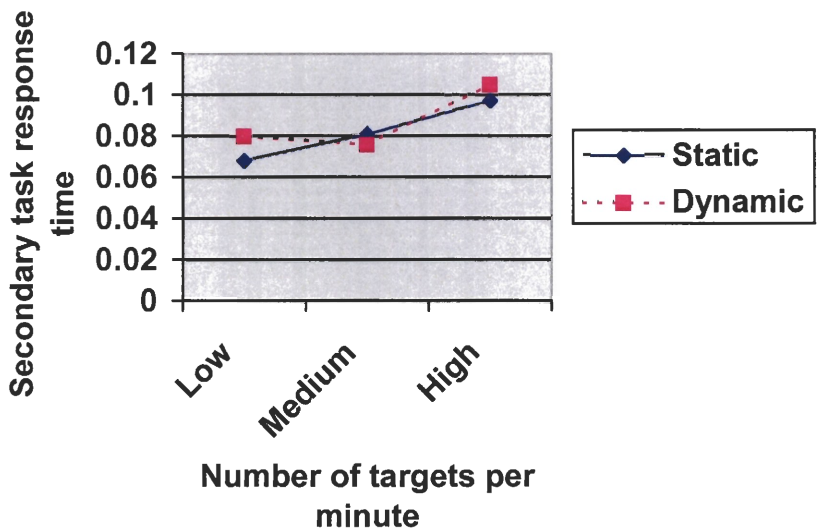

Figure 23. Adapting type by number of targets per minute for secondary task response time

Both attention theories predicted that performance in the static condition would be faster at the low number of targets per minute and performance in the dynamic would be faster at high numbers of targets per minute for the secondary task. Neither of these predictions was supported, although there was a non-significant indication that performance in the static condition was faster at the low number of targets per minute. In addition, both theories predicted a decline in the static mode performance, which was evidenced. They also predicted stable performance for the dynamic mode, which was supported from low to medium task demands.

Secondary task percent correct

There were no differences in percent correct performance for the secondary task between the static and dynamic function allocation conditions at medium number of targets per minute. However, at the low level of task demand, there was a significant difference between the two conditions, $t(59) = 2.240$, $p = .029$. This effect shows that performance in the dynamic condition was less accurate than the performance in the static condition. This was predicted by both attention theories. There was not a significant difference in percent correct performance for the secondary task at the high number of targets per minute, although the trend was for the dynamic mode to be more accurate than the static FA.

Within the dynamic FA, there was a significant difference in percent correct performance on the secondary task between the low and medium number of targets, $t(59) = 4.077$, $p < .001$, where the medium number of targets per minute resulted in significantly less accurate secondary task performance than the low. For the static function allocation, there were significant differences in secondary task percent correct between all three levels of targets per minute. For the difference between the low and medium numbers of targets per minute, performance in the low was significantly more accurate than performance in the medium, $t(59) = 6.551$, $p < .001$. Similarly, low numbers of targets per minute showed significantly more accurate secondary task performance than high, $t(59) = 8.182$, $p < .001$. Performance at the medium numbers of targets per minute was significantly more accurate for secondary task performance than high numbers of targets, $t(59) = 4.075$, $p < .001$ (see Figure 24).

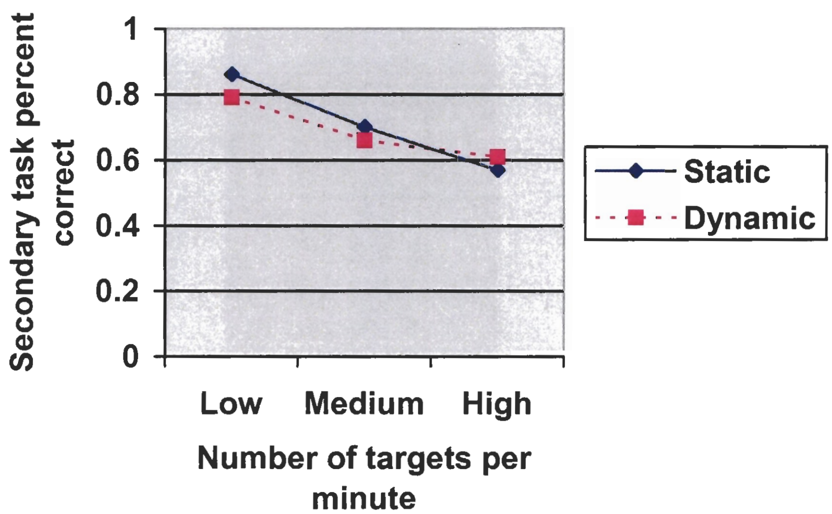

Figure 24. Adapting type by number of targets per minute for secondary task percent correct

Therefore, both theories were supported, in that performance declined in the static condition with increasing task demands. Performance in the dynamic condition was predicted to show no significant differences in secondary task percent correct throughout task performance, which was demonstrated between the medium and high number of targets per minute.

Examining Veto Mode versus Veto Mode

In addition to the comparison between the dynamic (with three levels of autonomy) and static conditions (with one level of autonomy), it is also interesting to examining the attention theories when comparing one level of autonomy each. This allowed a more direct comparison between the two function allocation conditions. When examining the static mode versus the dynamic function allocation when only in the veto mode, the results were quite interesting.

Primary task response time

For the response time performance in the primary task, there was a significant difference between participants in the static and dynamic function allocation conditions at low, $t(26) = -2.910$, $p = .007$, medium, $t(45) = -5.260$, $p < .001$, and high numbers of targets per minute, $t(39) = -3.919$, $p < .001$. In all cases, performance in the static condition was significantly faster than performance in the dynamic condition, as was predicted by both attention theories, although the high number of targets per minute was not expected to show significant differences.

As far as performance differences within one of the function allocation modes, there were significant differences in the response time performance on the primary task within the dynamic condition. Response time in the low number of targets was significantly faster than either performance in the medium, $t(24) = -3.292$, $p = .003$, or high number of targets per minute, $t(23) = 2.313$, $p = .030$. Similarly, within the static FA, there were significant differences in the primary task response time between the low number of targets and both the medium, $t(59) = -5.588$, $p < .001$, and high numbers of targets per minute, $t(59) = -5.780$, $p < .001$. In both cases, the participants interacting with the low number of targets per minute were significantly faster than when they interacted with either of the other two levels (see Figure 25).

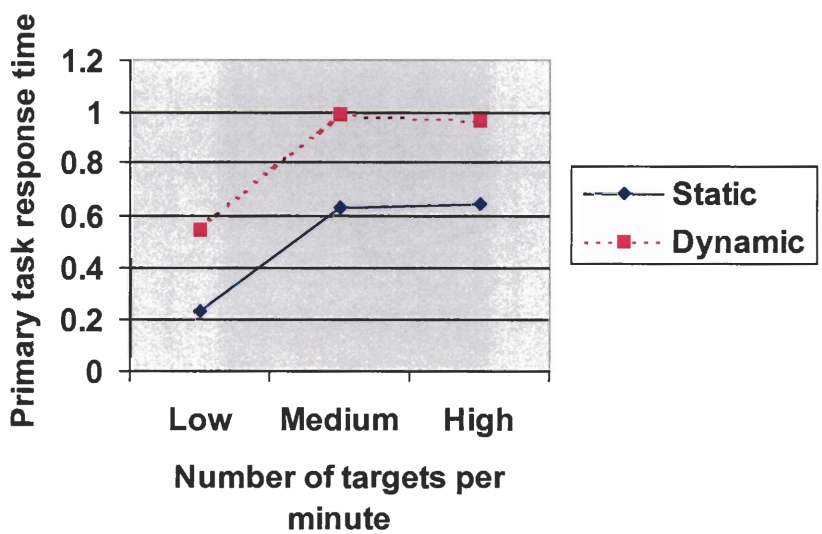

Figure 25. Adapting type by number of targets per minute for primary task response time (veto versus veto)

These functions are remarkably similar, with an apparent benefit in response time gained from being in the static condition. The results support MRT, which predicted a decline in the performance in the static FA (which occurred between low and medium numbers of targets) and no change in the performance in the dynamic FA (which occurred between medium and high numbers of targets per minute). Neumann's theory, on the other hand, was only partially supported by the static condition between medium and high number of targets per minute, where no significant changes in the response time performance for the primary task were noted.

Primary task percent correct

For percent correct in the primary task, there was only a significant difference between dynamic and static FA at the low, $t(25) = 2.483$, $p = .020$, and medium number of targets per minute, $t(45) = 2.439$, $p = .019$. In both of these cases, accuracy in the static condition was higher than in the dynamic condition. At the highest number of targets per minute, there was no difference in accuracy between the static and dynamic conditions, however, accuracy in the dynamic condition was slightly higher than in the static condition for the primary task. This supports both theories, as they predicted performance in the static condition would be superior at low demands, with performance differences between the conditions to be not significant at the high number of targets per minute.

There were no differences in primary task percent correct between the levels represented by the different number of targets per minute within the dynamic mode. For the static mode, there were significant differences between the high number of targets per minute and the low, $t(59) = 4.027$, $p <.001$, and medium ones, $t(59) = 3.376$, $p =.001$. In both cases, primary task accuracy in the high number of targets per minute condition was significantly lower than in either of the other two levels (see Figure 26).

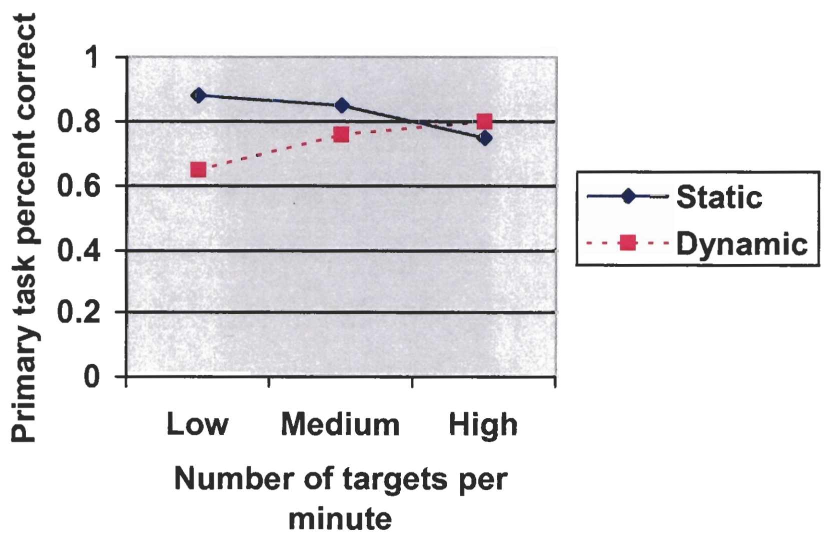

Figure 26. Adapting type by number of targets per minute for primary task percent correct (veto versus veto)

These findings were completely supportive of MRT, in that it was predicted that performance in the static condition would decline and performance in the dynamic condition would be stable. There was no support for Neumann's theory for the percent correct performance on the primary task.

Secondary task response time

There were no significant differences in secondary task response time between static and dynamic function allocation at any of the levels of numbers of targets per minute. Within the dynamic condition, there was a significant difference in response time performance on the secondary task between the medium and high number of targets, $t(33) = 3.636$, $p = .001$, with performance in high being slower than performance in the medium. The static condition resulted in significant response time differences between each of the levels. For the difference between the low and medium levels of static response time, the effect showed that low resulted in significantly slower secondary task response time than medium number of targets per minute, $t(59) = -4.046$, $p < .001$. Both low, $t(59) = -5.970$, $p < .001$, and medium, $t(59) = -4.484$, $p < .001$, resulted in significantly faster response times than the high numbers of targets per minute condition (see Figure 27).

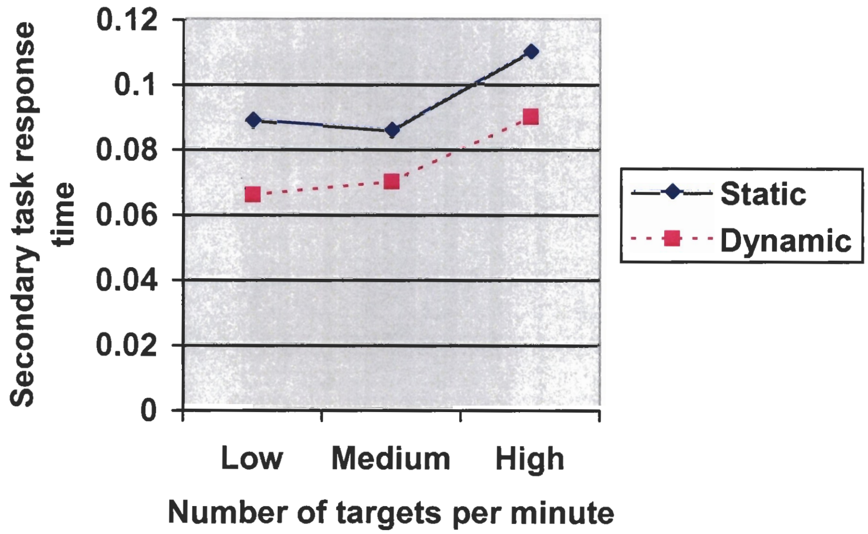

Figure 27. Adapting type by number of targets per minute for secondary task response time (veto versus veto)

Both attention theories found limited support from the response time data. The predictions were that the performance in the static condition would show a decline while performance in the dynamic condition would show stable secondary task response times. Reaction time in the static condition only showed a decline between medium and high numbers of targets, and reaction time in the dynamic condition was only stable between low and medium numbers of targets per minute, thus there was not overwhelming support for either of the theories.

Secondary task percent correct

There were no differences in accuracy between the static and dynamic function allocation conditions at either the low or high number of targets per minute. However, at the medium level, there was a significant difference in percent correct performance on the secondary task between the two conditions, $t(46) = 2.628$, $p = .012$. This effect showed that accuracy in the dynamic FA was lower than in the static FA. This was completely contrary to predictions, as the medium level was expected to show no differences in accuracy, whereas the other two levels were expected to show differences in accuracy.

Within the dynamic condition, there was a significant difference in the percent correct between the low and high number of targets per minute, $t(20) = 3.171$, $p = .005$. At the high number of targets per minute, percent correct on the secondary task was significantly worse than the percent correct at the low number of targets per minute. For the static function allocation, there were significant differences in accuracy between all three levels. For the difference between the low and medium numbers of targets per minute, secondary task accuracy in the low number of targets per minute condition was significantly higher than for the medium number of targets per minute condition, $t(59) = 6.551$, $p < .001$. Similarly, low numbers of targets per minute resulted in significantly higher accuracy than for the high number of targets per minute, $t(59) = 8.182$, $p < .001$. The medium number of targets was significantly more accurate for secondary task performance than the high numbers of targets per minute, $t(59) = 4.075$, $p < .001$ (see Figure 28).

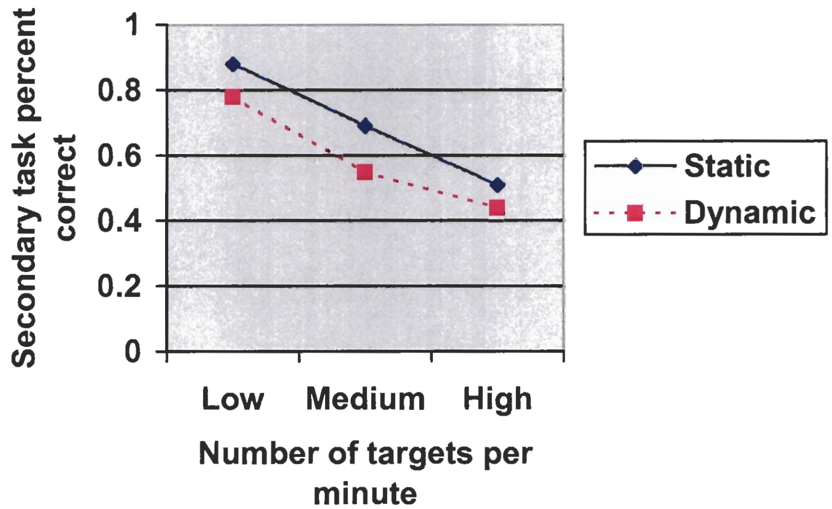

Figure 28. Adapting type by number of targets per minute for secondary task percent correct (veto versus veto)

In this case, the secondary task percent correct in the static condition declined as predicted by both theories. Accuracy in the dynamic condition, on the other hand, was predicted to be stable, which it was not. So, there was partial support for both theories, but only in the static condition.

To summarize the overall results relative to the predicted effects, the original predictions are shown below. Items in this table are bolded and underlined in areas where the results matched the predictions (see Table 4).

Table 4. Results versus Predicted Effects

	Response Time	Percent Correct	WL	SA
Adapting	Dynamic faster for secondary task- more attention available. **For primary task, dynamic slower than static until highest number of targets**	Dynamic more accurate for secondary task- more attention available. For primary task, dynamic less accurate than static until highest number of targets.	Dynamic lower - workload should be more constant with the changing allocations	Dynamic higher - more attention available to maintain SA
Number of targets per minute	**More targets - slower performance**	**More targets - less accurate performance**	More targets - higher workload	More targets - lower SA
Secondary Task	**If MRT, then ordering is dissimilar, then periphery & foveal** & if Neumann, foveal, periphery, dissimilar	If MRT, then ordering is dissimilar, then periphery & foveal & if Neumann, foveal, periphery, dissimilar	If MRT, then ordering is dissimilar, then periphery & foveal & if Neumann, foveal, periphery, dissimilar	**If MRT, then ordering is dissimilar, then periphery & foveal & if Neumann, foveal, periphery, dissimilar**
A X N	**Task more diff. -- dynamic the same, static gets slower (MRT) OR static the same and dynamic getting faster (Neumann) for primary task, secondary task - dynamic the same, static getting slower**	**Task more diff. -- dynamic the same, static gets less accurate (MRT)** OR static the same and dynamic gets more accurate (Neumann) for primary task, **secondary task - dynamic the same, static gets less accurate**	N/A	N/A
A X S	**No effect predicted - would indicate different reaction of secondary task (attention theory) when in different adapting mode**	No effect predicted - would indicate different reaction of secondary task (attention theory) when in different adapting mode	**No effect predicted - would indicate different reaction of secondary task (attention theory) when in different adapting mode**	No effect predicted - would indicate different reaction of secondary task (attention theory) when in different adapting mode
S X N	If MRT, interaction predicted, where tone less affected than foveal. **If Neumann- no interaction predicted**	If MRT, interaction predicted, where tone less affected than foveal. **If Neumann- no interaction predicted**	N/A	N/A
A X S X N	**No effect predicted - prediction would show difference in reactions of the secondary tasks (attention theories) when increasing difficulty in the static versus dynamic conditions**	No effect predicted - prediction would show difference in reactions of the secondary tasks (attention theories) when increasing difficulty in the static versus dynamic conditions	N/A	N/A

DISCUSSION

The major goals of this study were to a) investigate any benefits of dynamic function allocation (via a reduction in errors or performance improvements), b) examine which attention theory was supported using the secondary task methodology, and c) examine which attention theory was supported using the static versus dynamic automation manipulation. Based on these examinations, a number of measurement and theoretical issues were uncovered that are summarized in this section. Lastly, recommendations for designing automated systems, supporting mode awareness, and designing ground control stations are presented and discussed.

Benefits of dynamic function allocation (FA)

When examining the overall benefits of the dynamic FA using an ANOVA, there was very little support for an improvement in performance with the use of dynamic function allocation relative to static function allocation. For both primary task response time and primary task percent correct, there were no performance benefits from being in the dynamic FA. However, given the design was swayed to benefit performance in the static FA condition on these measures (as autonomous responses were not considered for the primary measures), it is perhaps not surprising that participants in the dynamic FA did not outperform the participants in the static FA for these measures.

Unfortunately, when examining other measures where the dynamic condition was predicted to result in better performance, it still did not, in most cases, show benefits relative to the static condition. For instance, performance in the dynamic condition was predicted to show improved situation awareness, workload, and mode awareness, given the additional attentional resources available to the participant. However, an examination of these results revealed that situation awareness and workload were not significantly better when participants experienced the dynamic condition. In fact, there were no differences in the situation awareness measures for static and dynamic FA, except for participants experiencing the periphery secondary task. These differences favored the static condition, which was contrary to predictions. There was no obvious reason for this difference, as participants with the periphery task did not experience more modes than those in the other two secondary tasks. They also did not report more mode awareness problems than other participants, nor did they report differences in the other situation awareness or workload scales.

There was not a direct way to measure differences in mode awareness, as participants in the static condition did not experience mode changes. However, mode awareness was examined for the dynamic FA. According to the exit interviews, two-thirds of the participants knew what mode they were in at all times. Thus, mode awareness did not appear to be a problem for the majority of participants. Although participants appeared to be aware of the mode they were in, this does not rule out the possibility that they may have an additional attention burden to maintain this awareness, thus reducing performance relative to the performance in the static FA.

The only positive finding with regards to dynamic function allocation benefits occurred when examining the covariance effect between secondary task response time and percent correct. This interaction revealed no speed-accuracy tradeoff, rather both the response time and percent correct declined as the number of targets per minute increased for participants in the static FA, but not for participants in the dynamic FA. This was consistent with predictions in that the participants in the dynamic condition had more resources available as the number of targets per minute increased, so the secondary task performance should not decline as rapidly.

Therefore, there was very limited support for the hypothesis that the dynamic FA would result in superior performance relative to the static one, particularly for performance on the primary task. The only benefit came in secondary task performance, suggesting future work may want to further examine the differing benefits of dynamic function allocation on the primary and secondary tasks. One possibility for the lack of findings on the primary task and in the workload and situation awareness measures was that the reduction in task demands inherent in the dynamic condition was offset by an increase in attention demands necessary to determine what mode was currently active. This could potentially be eliminated by further supporting mode awareness, although participants indicated they did not have difficulty with determining the active mode. Instead, it may be that this type of task was not conducive to a dynamic function allocation methodology because it fostered too many changes in the level of autonomy. Possibly a less dynamic task would benefit from dynamic FA, whereas this task was hindered by the number and speed of mode changes (about one every two minutes). However, given that this is an artificial environment augmented for experimental purposes, this does not rule out the possibility that dynamic FA could be successfully used in a real command and control task.

In addition, this study did not examine whether there were performance differences between static and dynamic function allocation as operators become more knowledgeable in the use of the system. Although novice operators, like those used in this study, did not benefit from the dynamic task allocation, it is possible more experienced operators might. A future study that trained operators to an expert level could explore this question. However, what is likely to occur is that experienced operators in the dynamic condition would spend an even larger majority of their time in the permission mode, while the operators in the static adapting type were in the veto mode. This would likely mitigate most of the benefits of being in the dynamic mode, as operators in the dynamic condition would be working harder than those in the static condition. Thus, this would require a redesign to the study to prevent that (e.g., have participants in the static FA be in permission mode instead of veto) and to provide a more equal comparison between the conditions.

Examining the Attention Theories with the Secondary Task Methodology

The primary comparison of the attention theories was accomplished using an ANOVA to determine if there were either main effects of the type of secondary task or an interaction between the type of secondary task and the number of targets per minute. There was only one significant main effect of secondary task, which occurred for secondary task response time. This effect was consistent with MRT, in that the tone task resulted in significantly better performance than either of the other two tasks.

With regards to the interaction effect, there was no significant interaction detected for primary task response time or secondary task response time and percent correct. This lack of a significant interaction is supportive of Neumann's view that the differences between the secondary tasks did not change with increasing task demands. There was a significant interaction for primary task percent correct, which would tend to support MRT. However, the nature of this interaction was opposite to that predicted by MRT; performance on the foveal task was the least affected by increasing the number of targets per minute and the tone task was most affected.

A possible explanation for these results is that the tone task was more "attention capturing" than the other two tasks, making it harder to ignore. When examining the secondary task performance, the tone task tended to result in better secondary task response time and percent correct (although no covariance was found between primary and secondary task performance). So, as the primary task became more demanding for participants in the tone task, the primary task performance declined, whereas the secondary task performance was maintained. This was not true for the other two secondary tasks. Thus, it is possible that the choice of this secondary task may have affected the results, although only for this dependent variable. Although this combination of secondary tasks was used previously, a future effort might benefit from a different choice of secondary tasks. However, it is important to note that this feature of the tone task may be beneficial in an actual system. For instance, if an alert message were auditory and verbal, similar to the tone task, it would be desirable that the message be attention capturing. Thus, this attention-capturing feature might actually be beneficial, depending on the usage within the system.

While the lack of interaction that occurred for the majority of variables would tend to support Neumann's theory, it is difficult to claim that Neumann's theory is supported when there was no significant main effect. Had there been main effects where the foveal task was clearly better than the periphery task and no interactions, it would have demonstrated clear evidence in favor of Neumann's view. Therefore, neither theory is strongly supported by this manipulation.

When examining these results to determine if they support the predictions made by the Proximity Compatibility Principle, it also did not garner support from these findings. The main effect was not consistent with PCP, as PCP was unable to distinguish between the tone task and the periphery task. PCP was also not supported by the interaction findings, as it also predicted an interaction between the secondary task and the number of targets per minute. Thus, none of the theories described in this paper garnered strong support from the secondary task manipulation.

Possible improvements to this manipulation for future work could be the use of different secondary tasks that were more demanding to further draw out potential performance differences between the tasks. However, given the current demands of the experimental task, it is unlikely that participants could have interacted with a more demanding secondary task. In addition, these "simple" secondary tasks were used in Schumacher, Seymour, Glass, Fenscik, Lauber, Kieras, & Meyer (2001), which was a less demanding experimental task, but still resulted in significant findings. Therefore, it is unlikely that the lack of significant findings is primarily due to the simplistic nature of the secondary tasks.

Another recommendation would be to reexamine the secondary tasks used, relative to the theories. MRT may have not been supported by the secondary task methodology because the primary task is not distinctive enough from the secondary tasks, particularly the dissimilar task. For instance, because both the primary and all of the secondary tasks required a decision followed by an execution with a mouse click, maybe they were not distinct enough to show experimental effects. Perhaps it would have been better to have the dissimilar (tone) task require a verbal response, instead of a mouse click. This might have enhanced the verbal/auditory versus spatial/visual differences between the tasks. If using these enhanced differences resulted in significant findings, this would indicate that MRT is supported when all of the processing codes are dissimilar, not just a subset. This would be an important finding for MRT, in that it clarifies when attention resources interfere with each other and when they should not interfere.

For Neumann's predictions, it is possible that the theory has to be somewhat refined in order find significant differences between the secondary tasks. For instance, Neumann's theory states that information close to the primary information can be attended to simultaneously with the primary information, however it may be this information must form an emergent feature with the primary information in order to be attended to, as is suggested by Allport (1989). Future work could examine the difference between simple proximity to major information and the development of emergent features to determine if Neumann's theory garners stronger support.

Examining the Attention Theories with Static versus Dynamic FA

When examining the attention theories and their predictions about performance with dynamic and static function allocation, there was mixed support for both theories, with most variables in the full data set showing partial support for both theories. In cases where Neumann's theory was supported, it tended to be at the higher task demands (between medium and high number of targets per minute), potentially suggesting that the low number of targets per minute was not sufficiently difficult to exhibit the predictions of Neumann's theory. To further examine this, a more difficult condition might provide further support for Neumann's view. This could be tested in a future study by adding one or more levels of task demand higher than those used in this experiment.

However, this explanation is somewhat counterintuitive, given the predictions of Neumann's theory. While it is possible to show a decline on secondary task performance in favor of improvements on primary task performance, there would some level of task demands where the primary task performance would begin to decline. Thus, it would seem more likely that Neumann's view would be more supported for the lower levels of task demands than the higher levels. While it is probable that the high task demands in this study were not above the "threshold" where primary task performance would decline, it is not clear why Neumann's view would not be supported at the low levels of task demand. So, although a future study could certainly examine the question, it seems somewhat unlikely that increasing the level of task demands would show additional support for Neumann's theory.

In addition to being supportive of Neumann's view, the higher levels of task demands could, at first glance, appear to support Malleable Attentional Resources Theory (MART). MART hypothesizes that available attention resources actually shrink to accommodate a reduction in task demand (Young & Stanton, 2002). Similarly, an increase in the size of the attention pool occurs up to a finite limit, where performance begins to fall off (Young & Stanton). The results from this study showed that, at the higher levels of task demands, static response time and percent correct were both stable, whereas the secondary task response time and percent correct both declined. This could indicate, for the static mode, that there were enough resources for the primary task to be stable but the secondary task had exceeded the threshold of task demand. While this appears to support MART, there would be much stronger evidence if the lower task demands (between low and medium number of targets per minute) showed stable performance. As is, the results indicate that the entire effect of MART (showing a trend of stable performance, followed by a decline) would have to appear between the medium and high number of targets per minute. If this were true, a future study with a fine grain analysis would be needed to examine the MART predictions more closely (i.e., using an examination of the attention ratio and the degree to which it changes based on task demands between the two levels utilized here). However, a research effect that is that specific seems to be somewhat unlikely, and does not appear to be a very probable explanation of the findings from this study.

At the lower task demands (between low and medium number of targets per minute), MRT was more often supported. This may suggest that MRT's predictions do not hold up well as the task demands increase beyond some criterion. This finding is actually supported by other research (see Wickens, 2000), where there was evidence that the dispersal of task processing across resources provided no benefit when the cognitive demands were high. Thus, above some criterion, there are not enough resources to be divided up between tasks, so the tenets of MRT begin to break down. Although this finding is supportive of MRT, it does not rule out other attention theories, like many of the single resource theories, where one pool of resources is not sufficient to support all of the task demands.

Time Spent in Each Level of Autonomy

One important improvement to this study would be to ensure that participants spent more time in the correct mode. The research predictions were based on participants being at higher levels of autonomy when they experienced higher task demands (in the dynamic mode), thus improving performance relative to the static mode. Although the pilot participants showed this, this study did not exhibit the same results (see Table 5). This table shows that the medium number of targets per minute resulted in performance in the veto mode 32% of the time. On the other hand, the high number of targets per minute resulted in more time in the permission mode (61% vs. 51%) and less time in the veto mode (19% vs. 32%) than the medium number of targets per minute. This may indicate that participants found the medium number of targets per minute more challenging than either of the other two levels, which is borne out when examining the data points for the medium level. For instance, in the dynamic condition, the medium level shows significantly worse performance than the other two levels for primary task response time and percent correct (when examining all of the data). However, for the secondary task, the higher level of autonomy (more time in the veto mode) tends to result in predicted trends in both response time and percent correct (i.e., steady declines as the number of targets per minute increase).

Table 5. Percentages of time spent in each level of autonomy

	Low	Medium	High
Permission	67	51	61
Veto	19	32	19
Autonomous	14	17	20

It is not clear why this inflated difficulty for the medium number of targets per minute occurred, but it could have been something artificial about the design (e.g., the secondary task had more events during the medium number of targets per minute or their location on the screen made them more difficult to find). However, this is unlikely, as the secondary task occurred at random intervals and the ordering of the targets per minute was changed for participants. In addition, nothing was apparent when the experiment was reexamined. Unfortunately though, this could have potentially affected the research theory predictions, possibly explaining why one of the theories was not strongly supported by the results.

A future direction of the study could be a redesign in which the percentage of time in the mode could be more directly controlled by the experimenter. Although this would eliminate the performance-based function allocation (by having the changes done based on the number of targets per minute instead of monitoring the participant's performance), it might elucidate the research theory findings in a more controlled manner. In addition, this redesign would provide a means to determine whether this inflated medium level of difficulty was a spurious result, as suspected, or a meaningful, replicable result. If meaningful, this finding would have interesting implications as to the selection of levels of autonomy in a dynamic FA system.

Implications for Attention Research

When trying to formulate implications for attention theories based on the overall findings described above, it is quite difficult. There is support for MRT at lower levels of task demand, followed by support for Neumann's theory at the higher levels of task demand. It is quite difficult to imagine a theory that would make predictions that performance would decline at low task demands, and then stabilizes as task demands increased. However, given that the medium levels of task demand appeared to be more challenging for participants (as evidenced by the higher percentages of time spent at higher levels of autonomy), it is not surprising that performance would stabilize as one went from medium to high levels of number of targets per minute. Thus, the overall results may not be the best data to utilize for examining the implications on attention research.

In order to discuss implications for attention research in the face of the above problem, there are two alternatives for looking at the data. The first would be to remove the medium number of targets data point and examine the trend of the other two points. When doing this, the static mode showed a significant decline in performance as task demands increased for both primary and secondary measures. This finding is consistent with MRT's predictions, where the increase in workload was expected to result in decreased performance. For the dynamic condition, there were mixed results, with one primary measure showing stable performance and the other showing a decline in performance. The same results were noted for the secondary measures. The dependent variables that showed stable performance in the dynamic FA would tend to support MRT, as increased task demands were offset by reductions in the workload. The declines in performance are somewhat harder to explain without a third data point. It could indicate no support for MRT (but also not for Neumann's view) or it could indicate that at the high level of task demands, there were not enough available resources even with the reduction in workload to maintain performance. In other words, the task was too demanding at high numbers of targets per minute. A follow-on study with a lower "high" data point or a redesign to fix the problem with medium targets would be very illuminating as to whether MRT is supported. Thus, this alternative approach to examining the data seems to provide support to MRT, with no support to Neumann's theory.

An alternative approach to reexamining the data would be to compare the static condition versus the dynamic condition when in the veto mode, as was described in the results section. In this comparison, only data when participants were in the veto mode were considered. So, although there may have been a larger percentage of the medium number of targets per minute data points in the comparison than might have been expected, there should be no problems with the time spent at each level of autonomy. Additionally, it is a more direct comparison between the static and dynamic modes. By using this comparison, it should alleviate the problems with the medium number of targets resulting in higher levels of autonomy than the high number of targets per minute.

When examining just these data, there was support for MRT for each dependent variable and almost no support for Neumann's theory. This was true both for the low to medium trend and the medium to high trend. For the static FA, the primary task variables showed declining performance as the demands increased. For the secondary task variables, there was evidence for a declining trend for the static FA throughout the levels of task demand. So, for the static condition, there was evidence that both primary and secondary task performance were affected by task demands, with increasing workload resulting in decreasing performance.

Performance in the dynamic FA was stable for the primary task measures and in secondary task measures during the lower task demands, as predicted by MRT. This would seem to indicate that the increasing levels of autonomy compensated for the increase in task demands. However, when examining secondary task performance during high task demands, performance declined, suggesting that the task had become too demanding (i.e., the workload was too high) for performance to be maintained. At this higher level, only secondary task performance declined, indicating that the participants were focusing more on the primary task, which they were instructed to do. An even higher level of task demand added to this study would likely have shown a greater decrement in the secondary task performance and/or a decrement in the primary task performance. Again, this examination seems to provide support for MRT, but not for Neumann's view.

Thus, the findings of both of these comparisons provide evidence in favor of MRT and the predictions about static versus dynamic function allocation based on MRT. These findings also support previous research about MRT, where above a threshold, there are not enough resources to be divided up between tasks, thus leading to a decline in performance and a breakdown on the tenets of MRT. Unfortunately, the secondary task manipulation , both using all the data or a subset, did not provide additional evidence to support MRT or further differentiate it from other theories. As mentioned previously, a replication of this study using different tasks may provide additional support for MRT, as it is possible that the secondary tasks selected interacted with the primary tasks in an unexpected way.

Measurement Issues

There is a related point to the attention research implications, as one examines the workload results. Much of the research in the previous section points to evidence supporting MRT, which would imply differences in the level of workload between the conditions. At the very least, one would expect to see an increase in the level of workload as the number of targets per minute increased for the static mode. However, all of the workload scales showed no significant effects. In the introduction, research was presented which claimed that this could occur if the automation was difficult to initiate or there were extensive data entry requirements (Parasuraman, Sheridan, & Wickens, 2000), neither of which were true in this study. Additionally, researchers have stated that primary task performance is more likely to show the impact of workload than secondary task performance (Scallen, Hancock, & Duley, 1995), which again was not supported by this research. Similar results were found when using the SART, with little noticeable effect being shown in the different scales of the SART, even though performance differences were evident.

Although there could be a theoretical underpinning to explain this lack of results, it seems more likely that the workload and SA measures used were simply not adequate to detect changes. Clearly, there were performance differences demonstrated, so it seems most likely that there were underlying differences in workload and SA that were not captured by the measures used. There are three potential reasons for why the measures used where not successful. The first is that the scales used where too large. It may be that participants are unable to meaningful discriminate between, for instance, 65 and 75 on a scale. Another potential problem with these measures is there was no behavior anchors associated with the scales. Because participants did not have a notion of what their responses should be, they may have been reluctant to indicate large changes in their workload or SA. In addition, they might have not been felt performance on this task was "hard", so they were not able to make the fine grained distinctions in performance that are required by these tools. The combination of the first and second problems suggest the employment of a new scale that has a smaller range with behavior anchors associated with each point on the scale (e.g., a 7-point Likert scale with performance descriptions associated with each point).

A third possible reason why these subjective tools did not correlate with performance differences is that participants were unaware of changes in their WL and SA. In other words, the subjective measure is not an effective way to measure either of these constructs, so one should utilize more objective measures of performance. For WL measurement, objective measures can include secondary task manipulations or physiological measurements. For SA, the SAGAT is considered an objective tool, as are methodologies when the task is stopped and questions are posed about the current situation. These, or other, methods may benefit the measurement of both WL and SA. However, perhaps the most valuable recommendation for the measurement of these constructs is to use multiple measures of WL and SA (e.g., a subjective and an objective measure) during the experimentation. This allows one to examine the correlation between the measures and allows the experimenter to more conclusively comment if one of the measures is not effective at measuring the construct. Thus, future work would likely benefit from a different measurement of workload and situation awareness or the use of different tools.

Theoretical Contributions

Because of the previously mentioned problem with the medium number of targets per minute, this section will discuss the subset of data only. The subset of data revealed much support for MRT, with no support for Neumann's theory. Although MRT was supported, it was not completely predictive of the results, so there are potentially improvements to be made to the theory.

First, there are issues with regards to the secondary task that must be addressed. Much of this was previously discussed in the secondary task discussion, where the issue of whether the dissimilar task was "too similar" was considered. If a task must be dissimilar on all three dimensions, processing code, stage, and modality, this should be made explicit in the literature. However, if this were the case, it would weaken the usefulness of MRT because in real world tasks, it is difficult to find jobs or tasks that are completely different from one another. More often, they are similar on one or two dimensions, and therefore would be expected to interfere more than a completely dissimilar task, but less than a completely similar task. In other words, some sort of linear function of interference would be expected. If, however, a completely dissimilar task (on all dimensions) were required for a significant effect, that would be an important finding for MRT's predictive power.

A second area that MRT might need to explore more is the suggestion from Neumann and others that differences in the proximity and location of information can affect performance. Although Neumann's predictions were not strongly supported in this area, there has been other research (i.e., movement time research, Keystroke Level Model, and the Goals, Operators, Methods, and Selection (GOMS) model) that has supported the notion that moving your eyes adds additional movement time to a task. It can also add attention demands if one has to remember a location to move to or search for the location. These findings, combined with research on emergent displays, point to implications for spatial location on task performance, which seem to be ignored in predictions made by MRT. By adding spatial location and movement time information to MRT based on the type of task (i.e., visual tasks have screen locations, whereas auditory tasks do not), one could increase the ability of MRT to predict operator performance.

Another interesting finding in this research was the decline in performance shown on the secondary task as the task demands increased (in the subset of data). As mentioned previously, this is a replication of previous findings were the tenets of MRT seem to break down as task demands are increased beyond some threshold. It would be interesting to examine this issue further to determine if this point can be manipulated or if the instructions can affect which task shows the decline in performance. For instance, a future study with two tasks that are equally weighted would be predicted to show a decline in both tasks, whereas emphasizing one of the tasks should show a decline in the other task as demands increase. This would be an interesting theoretical finding that could assist in task and system design, if one knew where dynamic function allocation could no longer compensate for increasing task demands. This point might represent where the system would have to switch to autonomous mode or where tasks would have to be reallocated to a different operator. This could have implications for teamwork research because the system could help decide who is currently overtasked and redistribute that operator's tasks to other operators in the system.

Design Recommendations

Attention Theories

Because both theories were supported to varying degrees in this experimental task when examining all of the data, the design recommendations actually consist of recommendation based on both theories. However, as is evident in the last section, the MRT design recommendations may be more applicable, given the support found for it during further evaluations. With regards to the maintenance of SA, MRT would predict that automation problems occur because too much attention and workload is directed at a different task (other than maintaining SA). Thus, a reduction of workload or attentional resources on other tasks would allow resources to be used for SA maintenance. This reduction was shown to be most helpful at the lower levels of task demands in this experimental task, suggesting a quicker switch to a higher level of autonomy may have been warranted.

On the other hand, the research about active processing would seem to support Neumann's view of attention, as active processing would suggest a participant was attending to the correct information and able to maintain their SA. If they were not focused correctly, they would lose SA. This awareness appears to be more crucial as the task demands increase, suggesting the need for better highlighting of information or increased volume/different frequency for the secondary task to be able to be recognized in an increasingly demanding environment.

Automated Systems

In addition to the attention theory design recommendations, there are also implications for the design of automation systems. First, this research showed that dynamic function allocation may not benefit the system as much as is desired or predicted by other research. Many current systems believe that adding dynamic function allocation will solve many of their problems and allow one operator to take on more tasks, however this study shows that performance in dynamic function allocation is not always better than performance in static FA. Before incorporating a dynamic function allocation system, a designer needs to consider the number of modes to be used, the maintenance of mode awareness, and the desired benefits of the system. In this study, performance in the dynamic FA was not better than performance in the static FA in most cases, possibly because the operators were switching modes too rapidly (about once every two minutes), which may be too fast for an operational system. Previous research that found benefits in the dynamic condition had only one mode switch when performance declined, with no switching when performance improved, thus may have failed to capture issues about speed and number of mode changes. Perhaps if too many modes are used or the switching is too rapid such that it becomes an additional burden for the operator, it may be beneficial to incorporate a static function allocation system. In either case, it is recommended that a system use operator-in-the-loop testing to ensure the desired benefits are derived from the system and display design chosen.

With regards to the use of an expert system to aid the operator in task performance, this study indicates that operators are able to interact, both statically and dynamically, with an expert system. Participants were not confused by the system's actions, nor did they have difficulty understanding how to interact with it. They also commented that they felt like they were in control of the system because they could always override the automation. This points to design recommendations for an expert system or any automation used to work collaboratively with humans. For successful interaction between humans and automation (either static or dynamic), one needs to ensure the system is predictable to the operator. This can include executing tasks in the same order as the human would, presenting displays in the order necessary for the human to carry out the steps, and using the same logic to reach a decision as the human. In addition, the automation should be as "transparent" as possible, so the human can see what factors led to a particular decision being made. This allows them to evaluate the correctness of the factors and ultimately the decision. This is particularly important when the human has received information that the automation is not privy to, making the automation's decision incorrect. Finally, operators need to have the ability to override the automation whenever necessary, enabling them to correct decisions that they determine are flawed.

Mode Awareness

One highlight of this study was in the participant's ability to maintain mode awareness. Although participants were not objectively measured on their mode awareness, their subjective comments indicated they knew what mode they were in at all times. Thus, a potential recommendation for the maintenance of mode awareness is to have a visual indicator of the current mode. Many current systems use a notification or alert of a mode change or simply a message in a communication log, all of which can be ignored or missed by an operator. By changing the visual appearance of the messages from the system, as was done in this experiment, or potentially using different colors or a different appearance to a header bar, a designer may be able to effectively support mode awareness, without requiring the operator to monitor a different part of the screen. Further study into these ideas, as well as objective measurement of their benefits, will likely improve future display designs for dynamic function allocation systems.

UCAV Ground Control Stations

Although this experiment did not provide a completely realistic environment for an UCAV operator, it does provide some insight into the design of an UCAV ground control station. For instance, the secondary tasks, while not actual tasks used by a future UCAV operator, do have similarities to collateral duties the operator would need to perform. The tone task is similar to a communications monitoring task of a UCAV operator, whereas the arrow tasks are similar to the alert and notification monitoring tasks that are currently being used. The results of this study, with its partial support for MRT, would indicate secondary tasks that are verbal and auditory might be preferable to visual, spatial alerts being used now. Using verbal communication for the communication task and tones for the alerts may reduce the demands on the visual channel and support better task performance by an operator. In addition, the use of 3-D audio to cue visual search efforts may also improve the primary task performance of operators, both in terms of accuracy and speed of target detection. Research that examines some of these design recommendations will likely advance the future designs of unmanned vehicle control stations.

Another implication of this study on ground control station design is the number of vehicles one operator can control. Many unmanned vehicle systems are considering having one operator control as many as ten vehicles at one time. This research used only four vehicles, albeit not in a completely realistic task environment. For a real system, current operator performance was probably not acceptable, particularly in percent correct performance, because a fielded system would be required to have minimal errors (~1%). The results from this study indicate that one operator controlling four vehicles would not be able to achieve this level of accuracy, at least initially. With more training, it is possible that one operator could successfully manage four vehicles with this level of automation, but that would have to be examined.

Conclusion

This research examined the benefits of dynamic function allocation, as well as examining the predictions of two attention theories on this adaptive automation. It also examined whether the attention theories, MRT and Neumann's theory, could explain the results found in the experiment, when utilizing a secondary task methodology. It did this by using a more complex task environment than previous studies, in this case a participant controlling multiple Unmanned Combat Air Vehicles (UCAVs), as well as three levels of autonomy, unlike many previous experiments that compared manual versus autonomous performance. While all of these manipulations had significant findings, none were strongly supportive of either of the research theories discussed in this paper. Subsequent examination with a subset of the data did provide more support for MRT in terms of primary and secondary task performance, but not in terms of workload or situation awareness.

All of the findings, however, did point to some areas that need further study and theoretical shortfalls that should be addressed in the future. From this research, design recommendations based on the attention theories were detailed for automated systems, supporting the maintenance of mode awareness, and UCAV ground control stations. These findings and design recommendations can be applied to all UCAV systems, as well as to other automated environments, possibly reducing the documented problems inherent in these domains. In addition, it could potentially be used as a selection tool in which the level of workload that a participant can manage before experiencing performance declines could be measured. This would be useful in high-stress environments (e.g., military or air traffic control), as successful operators could be identified based on their ability to withstand increases in workload.

APPENDIX A

Functional Breakdown for Experimental Task

Function (per vehicle)	**Who can carry out function? ***	**Task for DFA?**	**Static?**
1.0 Launch from carrier			
1.1 Power up using cart	Maintenance	No	
1.2 Radio contact			
1.2.1 Send communication from vehicle	Vehicle	No	
1.2.2 Send communication to vehicle	Operator	No	
1.3 Pre-taxi checks			
1.3.1 Subsystems checks	Maintenance	No	
1.3.2 Mission configuration verification	Vehicle	No	
1.3.1.1 Send configuration info from vehicle	Vehicle	No	
1.3.1.2 Send configuration info to vehicle	Operator	No	
1.3.1.3 Confirm the configuration info matches	Vehicle	No	
1.4 Taxi			
1.4.1 Taxi vehicle on deck	Maintenance	No	
1.5 Pre-takeoff checks			
1.5.1 Subsystems checks	Maintenance	No	
1.5.2 Flight systems checks	Maintenance	No	
1.6 Position on catapult			
1.6.1 Attach to launch bar	Maintenance	No	
1.6.2 Set vehicle weight	Maintenance	No	
1.6.3 Tension the catapult	Maintenance	No	
1.7 Launch consent			
1.7.1 Approve launch	Operator	No	
1.7.2 Signal to launch officer (via a light)	Vehicle	No	
1.8 Launch			
1.8.1 Signal approval to Catapult shooter	Maintenance	No	
1.8.2 Catapult shooter launches	Maintenance	No	
2.0 Departing the carrier area			
2.1 Airborne			
2.1.1 Climb after launch	Vehicle	No	
2.1.2 Achieve positive rate of climb	Vehicle	No	
2.1.3 Make airborne radio call	**Vehicle/Operator**	**Yes**	**Vehicle**
2.2 Angels 2.5 radio call	**Vehicle/Operator**	**Yes**	**Vehicle**
2.2.1 Continue climb through 2500 feet	Vehicle	No	
2.3 Turn to intercept 10 DME arc			
2.3.1 Continue climbing	Vehicle	No	
2.3.2 Make standard degree left turn	Vehicle	No	
2.4 Intercept 10 DME arc			
2.4.1 Continue climbing	Vehicle	No	
2.4.2 Stop turn	Vehicle	No	
2.5 Arcing radio call	**Vehicle/Operator**	**Yes**	**Vehicle**

2.6 Turn to intercept outbound radial			
2.6.1 Continue climbing	Vehicle	No	
2.6.2 Make standard degree right turn	Vehicle	No	
2.7 Outbound radio call	**Vehicle/Operator**	**Yes**	**Vehicle**
2.8 Maintain outbound radial			
2.6.1 Continue climbing	Vehicle	No	
3.0 Ingress			
3.1 Fly inbound route(s)			
3.1.1 Follow waypoints			
3.1.1.1 Adjust heading to maintain course	Vehicle	No	
3.1.1.2 Adjust course to compensate for wind	Vehicle	No	
3.1.2 Level off at 40,000 feet	Vehicle	No	
3.1.3 Accelerate to 480 knots	Vehicle	No	
3.2 Rendezvous into formation			
3.2.1 Choose rendezvous waypoint	Vehicle	No	
3.2.2 Choose formation type with offset for each vehicle	**Vehicle/Operator**	**Yes**	**Operator**
3.2.3 Adjust speed to achieve waypoint at offset	Vehicle	No	
3.2.4 Achieve waypoint heading			
3.2.4.1 Adjust heading to maintain course	Vehicle	No	
3.2.4.2 Adjust course to compensate for wind	Vehicle	No	
3.3 Change formation			
3.3.1 Choose formation type with offset for each vehicle	**Vehicle/Operator**	**Yes**	**Operator**
3.3.2 Choose rendezvous waypoint	Vehicle	No	
3.3.3 Adjust speed to achieve waypoint at offset	Vehicle	No	
3.3.4 Achieve waypoint heading			
3.3.4.1 Adjust heading to maintain course	Vehicle	No	
3.3.4.2 Adjust course to compensate for wind	Vehicle	No	
4.0 Mission area operations			
4.1 Enter Area of Operations	Vehicle	No	
4.2 Determine new route			
4.2.1 Break formation	Vehicle	No	
4.2.2 Fly pre-planned route to pre-planned targets	Vehicle	No	
4.3 Fly to pre-planned SAR image targets			
4.3.1 Follow waypoints to pre-planned images	Vehicle	No	
4.3.1.1 Adjust heading to maintain course	Vehicle	No	
4.3.1.2 Adjust course to compensate for wind	Vehicle	No	
4.3.2 Turn SAR on	**Vehicle/Operator**	**Yes**	**Vehicle**
4.3.3 Take SAR image	Vehicle	No	
4.3.4 Transmit SAR image	**Vehicle/Operator**	**Yes**	**Operator**
4.4 Detect pop-up target	Vehicle sensor	No	
4.5 Plan for pop-up target			
4.5.1 Determine target priority	Vehicle/Operator	No	Vehicle
4.5.2 Determine target-weapon pairing	Vehicle/Operator	No	Vehicle
4.5.3 Determine target-vehicle pairing	**Vehicle/Operator**	**Yes**	**Operator**
4.5.4 Plan route of vehicle to the new target	Vehicle	No	

4.6 Engage pop-up target			
4.6.1 Follow waypoints to target	Vehicle	No	
4.6.1.1 Adjust heading to maintain course	Vehicle	No	
4.6.1.2 Adjust course to compensate for wind	Vehicle	No	
4.6.2 Arm weapon	**Vehicle/Operator**	**Yes**	**Vehicle**
4.6.3 Transmit coordinates	Vehicle	No	
4.6.4 Launch weapon	**Vehicle/Operator**	**Yes**	**Vehicle**
4.7 Depart Area of Operations			
5.0 Egress			
5.1 Fly outbound route(s)			
5.1.1 Follow waypoints	Vehicle	No	
5.1.1.1 Adjust heading to maintain course	Vehicle	No	
5.1.1.2 Adjust course to compensate for wind	Vehicle	No	
5.1.2 Descend from 40,000 feet to marshal pattern height	Vehicle	No	
5.1.3 Maintain 480 knots	Vehicle	No	
5.2 Rendezvous into formation			
5.2.1 Choose rendezvous waypoint	Vehicle	No	
5.2.2 Choose formation type with offset for each vehicle	**Vehicle/Operator**	**Yes**	**Operator**
5.2.3 Adjust speed to achieve waypoint at offset	Vehicle	No	
5.2.4 Achieve waypoint heading			
5.2.4.1 Adjust heading to maintain course	Vehicle	No	
5.2.4.2 Adjust course to compensate for wind	Vehicle	No	
5.3 Change formation			
5.3.1 Choose formation type with offset for each vehicle	**Vehicle/Operator**	**Yes**	**Operator**
5.3.2 Choose rendezvous waypoint	Vehicle	No	
5.3.3 Adjust speed to achieve waypoint at offset	Vehicle	No	
5.3.4 Achieve waypoint heading			
5.3.4.1 Adjust heading to maintain course	Vehicle	No	
5.3.4.2 Adjust course to compensate for wind	Vehicle	No	
5.4 Break formation	Vehicle	No	
6.0 Marshaling in carrier area			
6.1 Enter marshal pattern using teardrop entry			
6.1.1 First vehicle enter at 6,000 feet	Vehicle	No	
6.1.2 Each subsequent vehicle enter at 1,000 foot altitude interval	Vehicle	No	
6.1.3 Fly 30 degrees off the outbound course for 2 minutes	Vehicle	No	
6.1.4 Fly standard rate left turn for 1 minute	Vehicle	No	
6.1.5 Fly 2 minute inbound leg	Vehicle	No	
6.2 Fly racetrack route			
6.2.1 Fly 2 minute outbound leg	Vehicle	No	
6.2.2 Fly standard rate left turn for 1 minute	Vehicle	No	
6.2.3 Fly 2 minute inbound leg	Vehicle	No	
6.2.4 Fly standard rate left turn for 1 minute	Vehicle	No	
6.3 Receive push time from ATC	Vehicle	No	

6.4 Adjust speed to meet push time	Vehicle	No	
6.5 Intercept 10 DME arc to land (left turn)	Vehicle	No	
6.6 Make marshal radio call	**Vehicle/Operator**	**Yes**	**Vehicle**
7.0 Landing on carrier			
7.1 Fly 10 DME arc	Vehicle	No	
7.2 Intercept Final Bearing			
7.2.1 Intercept extended runway bearing at 10 DME	Vehicle	No	
7.2.2 Begin descent to 140 knots	Vehicle	No	
7.3 Make Final Bearing radio call	**Vehicle/Operator**	**Yes**	**Vehicle**
7.4 Lower landing gear	Vehicle	No	
7.5 Lower tailhook	Vehicle	No	
7.6 Make Coupled with JPALS radio call	**Vehicle/Operator**	**Yes**	**Vehicle**
7.7 Follow descent route to carrier			
7.7.1 Adjust speed to maintain descent	Vehicle	No	
7.7.2 Make course corrections to compensate for carrier movement	Vehicle	No	
7.7.3 Adjust course to compensate for wind	Vehicle	No	
7.8 Make ball radio call	**Vehicle/Operator**	**Yes**	**Vehicle**
7.9 Touch down on deck	Vehicle	No	
7.10 Catch wire	Vehicle	No	
7.11 Taxi off runway	Maintenance	No	
* This breakdown is done for a vehicle capable of autonomous flight. The operator does not have the ability to control each vehicle manually because there are multiple vehicles and these inputs are considered flight critical (immediate). As such, the vehicle uses an autopilot to control course and subsystem function.			

APPENDIX B

NASA TLX

MENTAL DEMAND
How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

Low **High**

PHYSICAL DEMAND
How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Low **High**

TEMPORAL DEMAND
How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was that pace slow and leisurely or rapid and frantic?

Low **High**

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

Good **Poor**

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

Low **High**

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

Low **High**

APPENDIX C

Situation Awareness Rating Technique (SART)

Instructions:

SART uses subjective estimates of personal and task-dependent factors, which affect task performance and understanding to measure Situation Awareness in maneuver exercises. Situation Awareness refers to your ability to relate the meaning of events and elements in a noisy, uncertain environment to mission goals and objectives. The technique involves the scoring of fourteen different scales, each of which is potentially a factor in your Situation Awareness. This must be done just after debrief following performance on this task.

Remember the scales are a subjective measure of your individual perceptions during the specific exercise in the context of your experience with controlling and managing UCAVs in general. There is no right or wrong answer to give, only your best estimate of your personal experience. Do not spend too much time on any one item. Your initial 'gut feeling' as the exercise is likely the most accurate estimation.

The following are the definitions of each of the 14 SART rating items. Please read through these until you are sure you understand their meanings.

1. Demand on Cognitive Resources

How demanding was the exercise on your cognitive resources? Were there many difficult decisions and situations demanding constant attention and maximum efforts (high) or was it easy and minimally demanding (low)?

Low **High**

2. Instability of Situations

How changeable were the situations and environmental factors encountered through the course of the exercise? Were they very dynamic and likely to change suddenly (high), or were most of them slow and stable with easily predictable outcomes (low)?

Low **High**

3. Complexity of Situations

How complicated were the situations? Were they complex with many interrelated components and phases (high), or were most simple and straight forward (low)?

Low **High**

4. Variability of Situations

How many elements were changing at any one time in a given situation? Were there a large number of dynamic variables (high), or very few that might change at once (low)?

Low **High**

5. Supply of Cognitive Resources

How great a supply of cognitive and attentional resources coupled with decision aids and analysis tools did you have for problem-solving, decision-making, and other functions during the exercise? Could you bring a very large capacity to bear on the problems (high), or did you have limited resources at your disposal in each situation (low)?

Low **High**

6. Readiness

How alert and ready for action did you feel throughout the course of the exercise? Could you anticipate the flow of events and respond quickly (high), or were you hard pressed to keep up with evolving situations (low)?

Low **High**

7. Concentration of Attention

How much could you concentrate your attention in each problem situation? Were you always focused on important elements and events (high), or did technical details, controls, and displays distract you and draw your attention elsewhere (low)?

Low **High**

8. Division of Attention

Were you able to divide your attention among several key issues in the course of the exercise? Were you usually concerned with many aspects of current and future events simultaneously (high), or did you focus on only one thing at a time (low)?

Low **High**

9. Spare Mental Capacity

How much mental capacity did you have to spare in this exercise? Do you think you could have dealt with a significant number of additional elements and variables if necessary (high), or did the complexity of the exercise take all your mental capacity combined with available decision aids and analysis tools to handle (low)?

Low **High**

10. Understanding of the Situation

How well did you understand the tactical situations, the problems, and tasks presented in the exercise just run? Include ownship, all contacts, and all sources of information, as well as mission goals, strategy, and tactics for this purpose. In retrospect, did you usually have a good understanding in most cases (high), or did you have many unknowns and uncertainty a major part of the time (low)?

Low **High**

11. Information Quantity

How much useful information were you able to obtain from all available sources in the exercise? Did you receive and understand a great deal of pertinent data (high), or was very little of the information of much use for your task at hand (low)?

Low **High**

12. Information Quality

How good was the information you obtained about the situation? Was the knowledge communicated via all sources very accurate and precise (high), or was it noisy with high levels of uncertainty (low)?

Low **High**

13. Familiarity with Environment

How familiar were you with the different elements and events in the environment and situations encountered during the course of the exercise? Could you call on a great deal of relevant experience and knowledge to fill in gaps in the available information (high), or did you find significant aspects of the exercise new and unfamiliar to you (low)?

Low **High**

14. Situation Awareness

Evaluate your awareness of the overall meaning of events and elements in the environment to the mission plan and eventual accomplishment of mission goals. Did you always have a complete picture and a plan for how the various elements would affect the mission and could you anticipate future mission-critical events and decisions well in advance (high), or did you have very limited ability to predict the impact of on-going activity on future events and overall mission goals (low)?

Low **High**

APPENDIX D

Pilot Results

The question, regarding whether the two attentional theories are supported by the dynamic versus static automation manipulation, was examined using a t-test to determine if trends existed in the current data and if they were significant. As mentioned, the two attentional theories had different predictions as to the performance trends expected in the dynamic and static function allocation conditions. The results showed participants responded significantly faster to the targets in the low difficulty condition versus the high condition, when in the static mode ($t(5) = -2.651$, $p = .045$). Similarly, in the dynamic mode, participants responded significantly faster to the targets in the low difficulty condition than the medium difficulty ($t(5) = -3.424$, $p = .019$). There was also a significant difference between the modes at the high number of targets, with the static mode being significantly slower than the dynamic mode ($t(5) = 3.156$, $p = .025$). In this case, the static mode was getting slower, whereas the dynamic mode was closer to stable (see Figure 29). This would support MRT more, as dynamic was expected to be stable and static was supposed to decline.

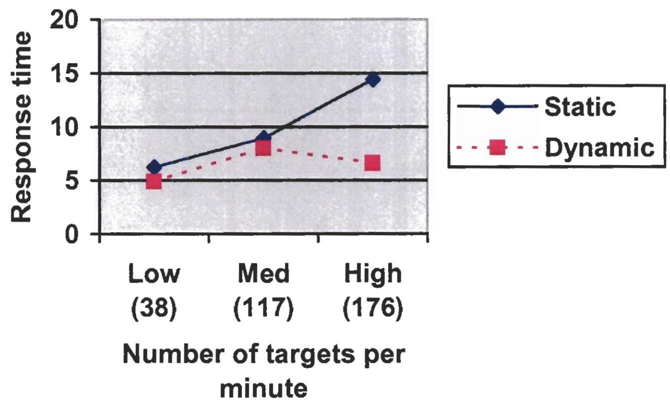

Figure 29. Pilot - Number of Targets per Minute by Adapting Type for Response time

There were no significant differences in the percent correct. However, the trend indicated that the static mode increased in percent correct, whereas the dynamic was closer to stable (see Figure 30). This was exactly opposite from the predictions of Neumann's theory, and only partial supportive of MRT. There did not appear to be a response time-accuracy tradeoff between the modes, as the static mode was slower and less accurate than the dynamic. There did, however, appear to be a speed-accuracy tradeoff within the static mode, as the participants get slower but more accurate as the task continued.

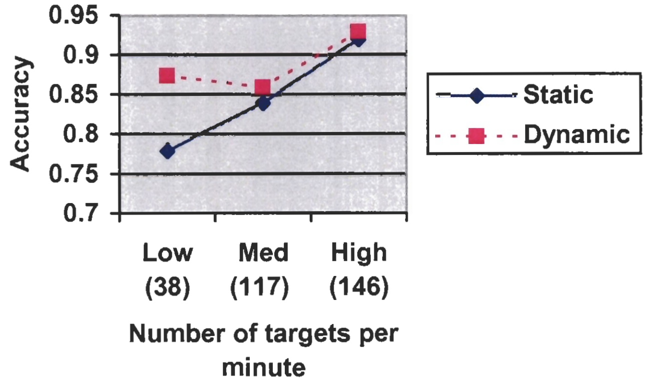

Figure 30. Number of Targets per Minute by Adapting Type for Accuracy

The second major hypothesis was that there would be differences between static and dynamic function allocation because it allowed increased interaction at low stress times and less interaction at high stress times, thereby stabilizing the participant performance and workload. This was examined using an ANOVA with workload, SA, response time, and percent correct as the dependent variables. This ANOVA revealed no significant effects for percent correct, workload, or SA and a significant effect of the mode for response time ($F(1,6) = 10.59$, $p = .047$). This revealed that the participants in the static mode took significantly longer than those in the dynamic mode. There was also a marginally significant effect of the number of targets for response time ($F(2, 6) = 4.592$, $p =.062$), in which the higher numbers of targets resulted in slower responses than the fewer targets. Contrasts of these values showed that the participants in the lower number of targets were significantly faster than they were in the medium ($t(5) = -3.267$, $p = .022$) or high ($t(5) = -2.946$, $p = .032$) numbers of targets. There were no differences between the secondary tasks.

APPENDIX E

ANOVA and MANOVA Tables

Table 6. PT Response Time

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	24.531	1	24.531	29.000	.000
Number of Targets per Minute	18.546	2	9.273	55.368	.000
Secondary Task	5.364	2	2.682	1.646	.202
Adapting Type by Secondary Task	6.340	2	3.170	3.748	.030
Number of Targets per Minute by Secondary Task	.300	4	.075	.447	.774
Adapting Type by Number of Targets Per Minute	3.732	2	1.866	18.502	.000
Adapting Type by Number of Targets per Minute by Secondary Task	.587	4	.147	1.455	.221
Total	92.872	57			

Table 7. PT Percent Correct

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	.0354	1	.0354	1.449	.234
Number of Targets per Minute	.472	2	.236	11.121	.000
Secondary Task	.013	2	.006	.314	.732
Adapting Type by Secondary Task	.100	2	.0502	2.056	.137
Number of Targets per Minute by Secondary Task	.388	4	.097	4.573	.002
Adapting Type by Number of Targets Per Minute	.222	2	.111	4.493	.013
Adapting Type by Number of Targets per Minute by Secondary Task	.098	4	.025	.992	.415
Total	1.185	57			

Table 8. ST Response Time

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	.013	1	.013	1.043	.311
Number of Targets per Minute	.039	2	.019	38.132	.000
Secondary Task	.085	2	.042	3.152	.050
Adapting Type by Secondary Task	.030	2	.015	1.197	.310

Number of Targets per Minute by Secondary Task	.001	4	.003	.596	.666
Adapting Type by Number of Targets Per Minute	.001	2	.000	1.450	.239
Adapting Type by Number of Targets per Minute by Secondary Task	.001	4	.0000	.908	.462
Total	.768	57			

Table 9. ST Percent Correct

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	.052	1	.052	1.248	.269
Number of Targets per Minute	3.543	2	1.772	54.719	.000
Secondary Task	1.030	2	.515	3.442	.039
Adapting Type by Secondary Task	.023	2	.011	.273	.762
Number of Targets per Minute by Secondary Task	.237	4	.060	1.827	.128
Adapting Type by Number of Targets Per Minute	.172	2	.086	2.451	.035
Adapting Type by Number of Targets per Minute by Secondary Task	.329	4	.082	3.306	.013
Total	8.529	57			

Table 10. Total Workload

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	120.00	1	120.00	.904	.346
Secondary Task	565.367	2	282.684	.897	.413
Adapting Type by Secondary Task	551.311	2	275.655	2.078	.135
Total	17953.377	57			

Table 11. Mental Demand

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	87.552	1	87.552	.338	.563
Secondary Task	1091.354	2	545.677	1.063	.352
Adapting Type by Secondary Task	1444.479	2	722.240	2.789	.070
Total	29247.969	57			

Table 12. Physical Demand

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	40.833	1	40.833	.270	.605
Secondary Task	114.687	2	57.344	.058	.944
Adapting Type by Secondary Task	797.604	2	398.802	2.636	.080
Total	56280.312	57			

Table 13. Temporal Demand

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	7.500	1	7.500	.030	.864
Secondary Task	1772.812	2	886.406	1.269	.289
Adapting Type by Secondary Task	312.187	2	156.094	.617	.543
Total	39806.562	57			

Table 14. Performance

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	977.552	1	977.552	3.489	.067
Secondary Task	3647.604	2	1832.802	2.874	.065
Adapting Type by Secondary Task	740.729	2	370.365	1.322	.275
Total	17953.377	57			

Table 15. Effort

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	220.052	1	220.552	1.106	.297
Secondary Task	180.000	2	90.000	.180	.835
Adapting Type by Secondary Task	232.917	2	116.458	.585	.560
Total	28442.906	57			

Table 16. Frustration

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	43.802	1	43.802	.210	.649

Secondary Task	2113.854	2	1056.927	1.139	.327
Adapting Type by Secondary Task	589.479	2	294.740	1.412	.252
Total	52870.469	57			

Table 17. Total SA

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	53.333	1	53.333	.293	.590
Secondary Task	216.354	2	108.177	.124	.883
Adapting Type by Secondary Task	158.853	2	79.427	.436	.649
Total	49560.938	57			

Table 18. Cognitive Demand

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	6.302	1	6.302	.027	.870
Secondary Task	1047.187	2	523.594	.826	.443
Adapting Type by Secondary Task	600.729	2	300.365	1.296	.282
Total	36145.469	57			

Table 19. Instability of Situations

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	46.875	1	46.875	.146	.704
Secondary Task	1556.354	2	778.177	1.399	.255
Adapting Type by Secondary Task	2509.063	2	1254.531	3.914	.026
Total	31697.812	57			

Table 20. Complexity of Situations

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	8.802	1	8.802	.023	.879
Secondary Task	2880.104	2	1440.052	1.743	.184
Adapting Type by Secondary Task	369.479	2	184.740	.487	.617
Total	47096.719	57			

Table 21. Variability of Situations

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	665.052	1	665.052	2.440	.124
Secondary Task	1298.229	2	649.115	1.036	.361
Adapting Type by Secondary Task	1886.354	2	943.177	3.460	.038
Total	35716.094	57			

Table 22. Supply of Resources

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	880.208	1	880.208	5.075	.028
Secondary Task	1466.354	2	733.177	.986	.380
Adapting Type by Secondary Task	821.979	2	410.990	2.370	.103
Total	42405.312	57			

Table 23. Readiness

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	110.208	1	110.208	.409	.525
Secondary Task	681.354	2	340.677	.395	.676
Adapting Type by Secondary Task	866.354	2	433.177	1.608	.209
Total	49189.687	57			

Table 24. Concentration of Attention

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	1.302	1	1.302	.008	.929
Secondary Task	907.917	2	453.958	.482	.620
Adapting Type by Secondary Task	291.667	2	145.833	.905	.410
Total	53655.156	57			

Table 25. Division of Attention

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	270.000	1	270.000	1.196	.279

Secondary Task	326.667	2	163.333	.224	.800
Adapting Type by Secondary Task	405.000	2	202.500	.897	.413
Total	41579.375	57			

Table 26. Spare Mental Capacity

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	206.719	1	206.719	.777	.382
Secondary Task	770.104	2	385.052	.486	.618
Adapting Type by Secondary Task	250.312	2	75.156	.282	.755
Total	45182.969	57			

Table 27. Understanding of Situations

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	35.208	1	35.208	.288	.594
Secondary Task	1173.229	2	586.615	.596	.554
Adapting Type by Secondary Task	441.979	2	220.990	1.808	.173
Total	56110.312	57			

Table 28. Quantity of Information

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	.208	1	.208	.001	.976
Secondary Task	83.750	2	41.875	.052	.950
Adapting Type by Secondary Task	395.417	2	197.708	.900	.412
Total	46235.000	57			

Table 29. Quality of Information

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	25.208	1	25.208	.094	.760
Secondary Task	738.229	2	369.115	.476	.623
Adapting Type by Secondary Task	1473.229	2	736.615	2.755	.072
Total	774.753	57			

Table 30. Familiarity of Situations

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	67.500	1	67.500	.305	.583
Secondary Task	3098.750	2	1549.375	1.435	.247
Adapting Type by Secondary Task	1235.000	2	617.500	2.793	.070
Total	61557.500	57			

Table 31. MANOVA table for PT Response Time

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	.235	1	.235	.882	.352
Secondary Task	1.687	2	.843	1.547	.222
PT Percent Correct	1.382	1	1.382	2.534	.117
ST Response Time	.307	1	.307	.563	.456
ST Percent Correct	.096	1	.096	.177	.675
Adapting Type by Secondary Task	1.864	2	.932	2.498	.070
Adapting Type by PT Percent Corr	.232	1	.232	.871	.355
Adapting Type by ST RT	.296	1	.296	1.112	.296
Adapting Type by ST Percent Corr	.658	1	.658	2.470	.122
Total	29.438	54			

Table 32. MANOVA table for PT Percent Correct

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	.019	1	.019	2.414	.126
Secondary Task	.007	2	.003	.526	.594
PT Response Time	.016	1	.016	2.534	.117
ST Response Time	.012	1	.012	1.826	.182
ST Percent Correct	.024	1	.024	3.720	.059
Adapting Type by Secondary Task	.023	2	.011	1.459	.242
Adapting Type by PT RT	.003	1	.003	.400	.530
Adapting Type by ST RT	.000	1	.000	.014	.907
Adapting Type by ST Percent Corr	.028	1	.028	3.577	.064
Total	.350	54			

Table 33. MANOVA table for ST Response Time

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	.011	1	.011	1.102	.298
Secondary Task	.035	2	.017	2.037	.140
PT Response Time	.004	1	.004	.563	.456
PT Percent Correct	.015	1	.015	1.826	.182
ST Percent Correct	.039	1	.039	4.605	.036
Adapting Type by Secondary Task	.004	2	.002	.221	.802
Adapting Type by PT RT	.016	1	.016	1.653	.204
Adapting Type by PT Percent Corr	.011	1	.011	1.129	.293
Adapting Type by ST Percent Corr	.000	1	.000	.012	.913
Total	.467	54			

Table 34. MANOVA table for ST Percent Correct

Source	Sum of Squares	DOF	Mean Square	F	Sig.
Adapting Type	.321	1	.321	4.387	.061
Secondary Task	.077	2	.038	.398	.673
PT Response Time	.017	1	.017	.177	.675
PT Percent Correct	.359	1	.359	3.720	.059
ST Response Time	.444	1	.444	4.605	.036
Adapting Type by Secondary Task	.105	2	.052	.718	.492
Adapting Type by PT RT	.001	1	.001	.014	.906
Adapting Type by PT Percent Corr	.344	1	.344	4.706	.064
Adapting Type by ST RT	.000	1	.000	.006	.939
Total	5.207	54			

REFERENCES

Allport, A. (1989). Visual Attention. In M. I. Posner (Ed.), Foundations of Cognitive Science (pp. 631-682). Cambridge, MA: The MIT Press.

Allport, D. A. (1980). Attention and Performance. In G. Claxton (Ed.), Cognitive Psychology – new directions. London: Routledge and Kegen Paul.

Andes, R. C., Jr. (1990). Adaptive aiding automation for system control: Challenges to realization. Proceedings of the Topical Meeting on Advances in Human Factors Research on Man-Computer Interactions: Nuclear and Beyond, USA.

Bennett, K. B., Cress, J. D., Hettinger, L. J., Stautberg, D., & Haas, M. W. (2001). A theoretical analysis and preliminary investigation of dynamically adaptive interfaces. International Journal of Aviation Psychology, 11(2), 169-195.

Campbell, R. H. (2001). The Design of Displays to Support Decision Making in Procedural and Automated Domains. Unpublished Preliminary Examination, Georgia Institute of Technology.

Coury, B. G. & Boulette, M. D. (1992). Time stress and the processing of visual displays. Human Factors, 34(6), 707-725.

Degani, A., Shafto, M., & Kirlik, A. (1999). Modes in human-machine systems: Constructs, representation, and classification. The International Journal of Aviation Psychology, 9(2), 125-138.

Endsley, M. R. (1993). Situation awareness and workload: Flip sides of the same coin. Proceedings of the Seventh International Symposium of Aviation Psychology, USA, 906-911.

Endsley, M. R. (1996). Automation and Situation Awareness. In R. Parasuraman & M. Mouloua (Eds.), Automation and Human Performance (pp. 163-181). Mahwah, NJ: Lawrence Erlbaum Associates.

Endsley, M. R., Farley, T. C., Jones, W. M., Midkiff, A. H., & Hansman, R. J. (1998). Situation Awareness Information Requirements for Commercial Airline Pilots (Technical Report ICAT-98-1). Cambridge, MA: Massachusetts Institute of Technology, International Center for Air Transportation.

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

Fitts, P. M. (1951). Human Engineering for an effective air navigation and traffic control system. Washington, DC: National Research Council.

Galotti, K. M. (1998). Cognitive Psychology In and Out of the Laboratory. Belmont, CA: Wadsworth Publishing Company.

Greenstein & Revesman (1986). Two simulation studies investigating means of human-computer communication for dynamic task allocation. IEEE Transactions on Systems, Man, and Cybernetics, 16(5), 726-730.

Haas, M. W. & Hettinger, L. J. (2001). Current research in adaptive interfaces. The International Journal of Aviation Psychology, 11(2), 119-122.

Hancock, P. A. & Scallen S. F. (1998). Allocating functions in human-machine systems. In R. R. Hoffman & Sherrick, M. F. (Eds.), Viewing Psychology as a Whole: The Integrative Science of William N. Dember (pp. 509-539). Washington: American Psychological Association.

Haskell, I. D. & Wickens, C. D. (1993). Two- and three-dimensional displays for aviation: A theoretical and empirical comparison. International Journal of Aviation Psychology, 3(2), 87-109.

Hollnagel, E. & Woods, D. D. (1999). Cognitive systems engineering: New wine in new bottles. International Journal of Human-Computer Studies, 51, 339-356.

Javaux, D. (1998a). An algorithmic method for predicting pilot-mode interaction difficulties. Proceedings of DASC '98, Seattle, USA, 906-911.

Javaux, D. (1998b). Explaining Sarter & Woods' classical results: The cognitive complexity of pilot-autopilot interaction on the Boeing 737-EFIS. Proceedings of the 2nd Workshop on Human Error, Safety, and Systems Development, Seattle, USA, 62-77.

Javaux, D. & De Keyser, V. (1998). The cognitive complexity of pilot-mode interaction: A possible explanation of Sarter & Woods' classical results. Proceedings of the International Conference on Human-Computer Interaction in Aeronautics (HCI-Aero '98), Montreal, Canada.

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

Kantowitz, B. H. & Sorkin, R. D. (1987). Allocation of functions. In G. Salvendy (Ed.), Handbook of Human Factors (pp.355-369). New York: Wiley.

Kerr, B. (1973). Processing demands during mental operations. Memory and Cognition, 1, 401-412.

Luczak, H. (1997). Task Analysis. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics, (2nd ed., 340-416). New York: John Wiley & Sons.

Meyer, D. E. & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. Psychological Review, 104(1), 3-65.

Moray, N. (1967). Where is capacity limited: A survey and a model. Acta Psychologica, 27, 84-92.

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

Navon, D. & Gopher, D. (1979). On the economy of the human processing system. Psychological Review, 86, 254-255.

Neisser, U. (1976). Cognition and Reality. San Francisco: Freeman.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), Perspectives on Perception and Action (pp. 361-394). Hillsdale, NJ: Lawrence Erlbaum Associates.

Olson, W. A. & Sarter, N. B. (2001). Management by consent in human-machine systems: When and why it breaks down. Human Factors, 43(2), 255-266.

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

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

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

Raby, M & Wickens, C. D. (1994). Strategic workload management and decision biases in aviation. The International Journal of Aviation Psychology, 4(3), 211-240.

Rogers, W. A., Rousseau, G. K., & Fisk, A. D. (1999). Applications of Attention Research. In F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumas, D. S. Lindsey, & M. T. H. Chi (Eds.), Handbook of Applied Cognition (pp. 33-55). New York: John Wiley & Sons.

Rudolph, F. M. (2000). Human performance during automation: the interaction between automation, system information, and information display in a simulated flying task. Unpublished Dissertation. Georgia Institute of Technology.

Sanders, M. S. & McCormick, E. J. (1987). Human Factors in Engineering and Design. New York: McGraw-Hill.

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

Scallen, S. F. & Hancock, P. A. (2001). Implementing dynamic function allocation. International Journal of Aviation Psychology, 11(2), 197-221.

Scallen, S. F., Hancock, P. A., & Duley, J. A. (1995). Pilot performance and preference for short cycles of automation in adaptive function allocation. Applied Ergonomics, 26(6), 397-403.

Scerbo, M. W., Freeman, F. G., & Mikulka, P. J. (2000). A biocybernetic system for adaptive automation. In R. W. Backs & W. Boucsein (Eds.), Engineering Psychophysiology: Issues and Applications (pp. 241-253). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schneider, W. & Fisk, A. D. (1982). Concurrent automatic and controlled visual search: Can processing occur without resource cost? Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 261-278.

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. Psychological Science, 12(2), 101-108.

Selcon, S. J. & Taylor, R. M. (1996). Evaluation of the Situational Awareness Rating Technique (SART) as a tool for aircrew systems design. In Situational Awareness in Aerospace Operations, (AGARD-CP-478). (pp. 5-1 - 5-8). Neuilly Sur Seine, France: NATO AGARD.

Sheridan, T. B. & Verplank, W. L. (1978). Human and computer control of undersea teleoperators (Tech. Report). Man-Machine Systems Laboratory, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Strayer, D. L. & Kramer, A. F. (1994). Strategies and automaticity: I. Basic findings and conceptual framework. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20(2), 318-341.

Svensson, E. A. I. & Wilson, G. F. (2002). Psychological and psychophysiological models of pilot performance for systems development and mission evaluation. The International Journal of Aviation Psychology, 12(1), 95-110.

Tattersall, A. J. & Morgan, C. A. (1997). The function and effectiveness of dynamic task allocation. In D. Harris (Ed.), Engineering Psychology and Cognitive Ergonomics, Vol. 2: Job Design and Product Design. (247-255). Burlington, VT: Ashgate Publishing Co.

Taylor, R. M. (1990). Situation Awareness Rating Technique (SART): The development of a tool for aircrew systems design. In Situational Awareness in Aerospace Operations, (AGARD-CP-478). (pp. 3-1 - 3-17). Neuilly Sur Seine, France: NATO AGARD.

Taylor, R. M., Shadrake, R., Haugh, J. & Bunting, A. (1996). Situational awareness, trust, and compatibility: Using cognitive mapping techniques to investigate the relationship between important cognitive system variables. In Situation Awareness: Limitations and Enhancements in the Aviation Environment (AGARD-CP-575). (pp. 6-1 - 6-14). Neuilly Sur Seine, France: NATO AGARD.

Tsang, P. & Wilson, G. F. (1997). Mental workload. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics (2nd ed., pp. 418-449). New York: John Wiley & Sons.

Vidulich, M. A., McCoy, A. L., & Crabtree, M. S. (1996). Attentional control and situational awareness in complex air combat simulation. In Situation Awareness: Limitations and Enhancements in the Aviation Environment (AGARD-CP-575). (pp. 18-1 - 18-5). Neuilly Sur Seine, France: NATO AGARD.

Warm, J. S. & Dember, W. N. (1998). Tests of vigilance taxonomy. In R. R. Hoffman & Sherrick, M. F. (Eds.), Viewing Psychology as a Whole: The Integrative Science of William N. Dember (pp. 87-112). Washington: American Psychological Association.

Wickens, C. D. (2000). Imperfect and unreliable automation and its implications for attention allocation, information access, and situation awareness. (Technical Report No. ARL-00-10/NASA-00-2). Moffett Field, CA: NASA Ames Research Center.

Wickens, C. D (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), Varieties of Attention (pp.63-102). Orlando, FL: Academic Press.

Wickens, C. D. & Carswell, C. M. (1997). Information Processing. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics (2nd ed., pp. 89-129). New York: John Wiley & Sons.

Wickens, C. D., Gordon, S. E., & Liu, Y. (1997). An Introduction to Human Factors Engineering. New York: Addison Wesley Longman, Inc.

Wiener, E.L. & Curry, R. E. (1980). Flight-deck automation: Promises and problems. Ergonomics, 23(10), 995-1011.

Woods, D. D. (1996). Decomposing automation: Apparent simplicity, real complexity. In R. Parasuraman & M. Mouloua (Eds.), Automation and Human Performance (pp. 3-17). Mahwah, NJ: Lawrence Erlbaum Associates.

Young, M. S. & Stanton, N. A. (2002). Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. Human Factors, 44(3), 365-375.