2002

Individual differences and situation awareness

Jaruwan Klamklay
Iowa State University

Recommended Citation

Klamklay, Jaruwan, "Individual differences and situation awareness " (2002). *Retrospective Theses and Dissertations*. 1004.
https://lib.dr.iastate.edu/rtd/1004

Individual differences and situation awareness

by

Jaruwan Klamklay

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Industrial Engineering

Program of Study Committee:
Patrick E Patterson, Major Professor
Thomas Barta
Douglas Gemmill
Kenneth Koehler
Judy Vance

Iowa State University

Ames, Iowa

2002

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Graduate College

Iowa State University

This is to certify that the doctoral dissertation of

Jaruwan Klamklay

has met the dissertation requirements of Iowa State University

Major Professor

For the Major Program

Table of Contents

List of Figures

Chapter 2

Chapter 4

List of Tables

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Acknowledgments

I would not have been able to finish my dissertation if I did not have the support of friends and colleagues. I appreciated the opportunity to work with Dr. Patrick E Patterson, and would like to give my sincere thanks to him for his cooperation and supervision with this project. Thanks also to Dr. Thomas Barta, Dr. Douglas Gemmill, Dr. Kenneth Koehler, and Dr. Judy Vance for their time and input into my dissertation. Thanks to all my subjects for giving their time to participate in my experiment. Thanks to Justin Recknor and Dongmei Zhai, graduate students from the Iowa State University Statistics department, for the help with the data analysis.

Abstract

This study is divided into five sections, each with an objective to: (1) examine the attention and memory limitations of engineering students; (2) examine the effects of automation, pace, and task duration on situation awareness and task performance of novice operators of a complex system; (3) examine the effects of attention limitations on situations awareness and task performance; (4) examine the effects of memory limitations on situation awareness and task performance; and (5) identify the effects of individual differences on the relationships among workload, situation awareness, and task performance, respectively.

Eighty undergraduate engineering students performed nine psychological tests to measure individual differences in terms of attention and memory limitations. Speed and/or accuracy were used to score each test. The participants also performed a simplified pilot task, in which response accuracy and latency were used to measure operator performance, and the NASA-TLX to measure workload. The Situation Awareness Global Assessment Technique (SAGAT) and a bi-polar subjective rating scale were used as the measures of situation awareness.

The results showed that the attention and memory limitations of engineering students who participated in this study were similar to that of non-engineer participants of previous studies, however engineering students seemed to be less impacted by such limitations. Automation reduced situation awareness and yielded mixed effects on task performance. High task pace improved situation awareness and had mixed impacts on task performance. Task duration yielded mixed effects on task performance and had no impact on situation awareness. Operators with high scores on the attention tests and the memory tests tended to have high task performance and high situation awareness. When compared with the entire participant pool, groups of individuals with similar attention and memory limitations produced not only stronger correlation coefficients amongst workload, task performance, and situation awareness, but also produced different sets of correlated components.

Therefore, both system factors and individual factors must be taken into account in the early stages of system design and evaluation. Psychological tests may be used to identify individuals potentially having high performance and high situation awareness when working in a complex system.

Chapter 1. General Introduction

Introduction

Situation awareness (SA) has become a topic of ergonomic interest in recent years. Research studies have revealed that the development of situation awareness is highly related to operator performance. Generally, higher situation awareness is found to contribute to superior performance, especially for operators of dynamic complex systems, such as pilots (combat and non-combat), air traffic controllers, and manufacturing plant operators.

Studies have focused on two groups of factors affecting situation awareness: task/system factors, and individual factors. System factors such as automation, interface, and workload, were found to affect SA, and the findings have been reasonably consistent. Although improving system design has been found to be effective, limitations of space, cost, time, work process, and especially technologies are encountered in the application of this approach. Thus, it is almost impossible to eliminate every potential error by just improving the task/system factors, indicating the need to consider individual factors simultaneously.

In addition to improving system factors, the training and recruiting of individuals having special abilities for acquiring and maintaining situation awareness is becoming another attractive alternative. However, the findings of studies examining the affect of individual factors on situation awareness have been mixed. Some have concluded that an individual's attention and perception characteristics are good situation awareness predictors, while others have concluded that memory is a critical factor. In addition, some have found that computer game experience can improve SA, while others have found the opposite.

These inconsistent findings have resulted from factors such as inconsistencies in identifying individual differences, the use of small sample groups, and the SA measurement tools used.

Factors that could best describe individual differences might be the limitations of attention and memory. Attention and memory are crucial elements to information processing, a process generally used to describe operator performance and situation awareness.

In addition to the factors affecting SA, studies also focus on the relationships among operator workload, performance, and situation awareness which are the principle approaches used in design evaluation. To completely evaluate a system, parallel findings from all three approaches are needed to ensure an optimal design. However, subjective methods generally used to measure workload could be affected by individual differences, resulting in an invalid conclusion.

Therefore, the objectives of this study were to: (1) examine the attention and memory limitations of engineering students, (2) examine the effects of system factors on situation awareness and task performance, (3) examine the effects of limitations of human memory on situation awareness and task performance, (4) examine the effects of limitations of human attention on situation awareness and task performance, and (5) identify the effects of human limitations on the relationships between workload, situation awareness, and task performance.

A clearer understanding of the limitations of human attention and memory, and how task factors and individual factors affect operator situation awareness and performance would provide additional insight into the cause of accidents, designing controls and displays to support high SA, designing systematic training programs, or recruiting operators potentially to have high SA.

A firm knowledge of the relationships between workload, performance, and situation awareness could reduce the necessity of considering all three approaches during system evaluation. It could be applied to system evaluation strategies to decrease the cost and time, yet keep the reliability of the system evaluation at an acceptable level.

Dissertation Organization

This dissertation is organized as follows: Chapter Two covers a summary of situation awareness theory and related studies. It describes situation awareness (SA) and provides an SA model, its importance, measurement techniques, factors affecting SA, and previous related studies. Chapter Three presents the attention and memory limitations of engineering students. Chapter Four presents the effects of system factors on operator performance and situation awareness. Chapter Five presents the effects of attention limitations on operator performance and situation awareness. Chapter Six presents the effects of memory limitations on operator performance and situation awareness. Chapter Seven presents the relationships among perceived workload, task performance, and situation awareness with and without taking the attention and memory limitations into consideration. Due to an objective to divide this study into several individual papers, some information in chapters three to seven may be repeated. Finally, Chapter Eight presents conclusions and future work.

Chapter 2. Literature Review

A summary of situation awareness theory and related studies is presented in the following. The first section describes the meaning of situation awareness. This is then followed by it model, importance, and measurement techniques. The next section describes factors affecting situation awareness. Limitations of attention and memory, individual factors that could affect situation awareness, are then presented.

Definition of Situation Awareness

Situation awareness is the knowledge of the current state of a system. It is the understanding of what is going on in the system at a certain point of time, and what will happen in the near future. Endsley (in Endsley and Gerland, 2000) provided a simple definition of situation awareness "knowing what is going on around you".

Several definitions of situation awareness can be found in the literature (Judge, 1992; Vidulich, 1992; Endsley, 1988, 1995; Gugerty, 1998; Sarter and Woods,1995). Due to the uniqueness of each environment, the concept of SA was perceived differently across the different system domains. To prevent any possible confusion about the meaning of SA, this study adopts the concept defined by Endsley (1988), and Endsley (1995a). This definition of SA is applicable to wide variety of domains.

"the perception of 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" Endsly (1988 p87)

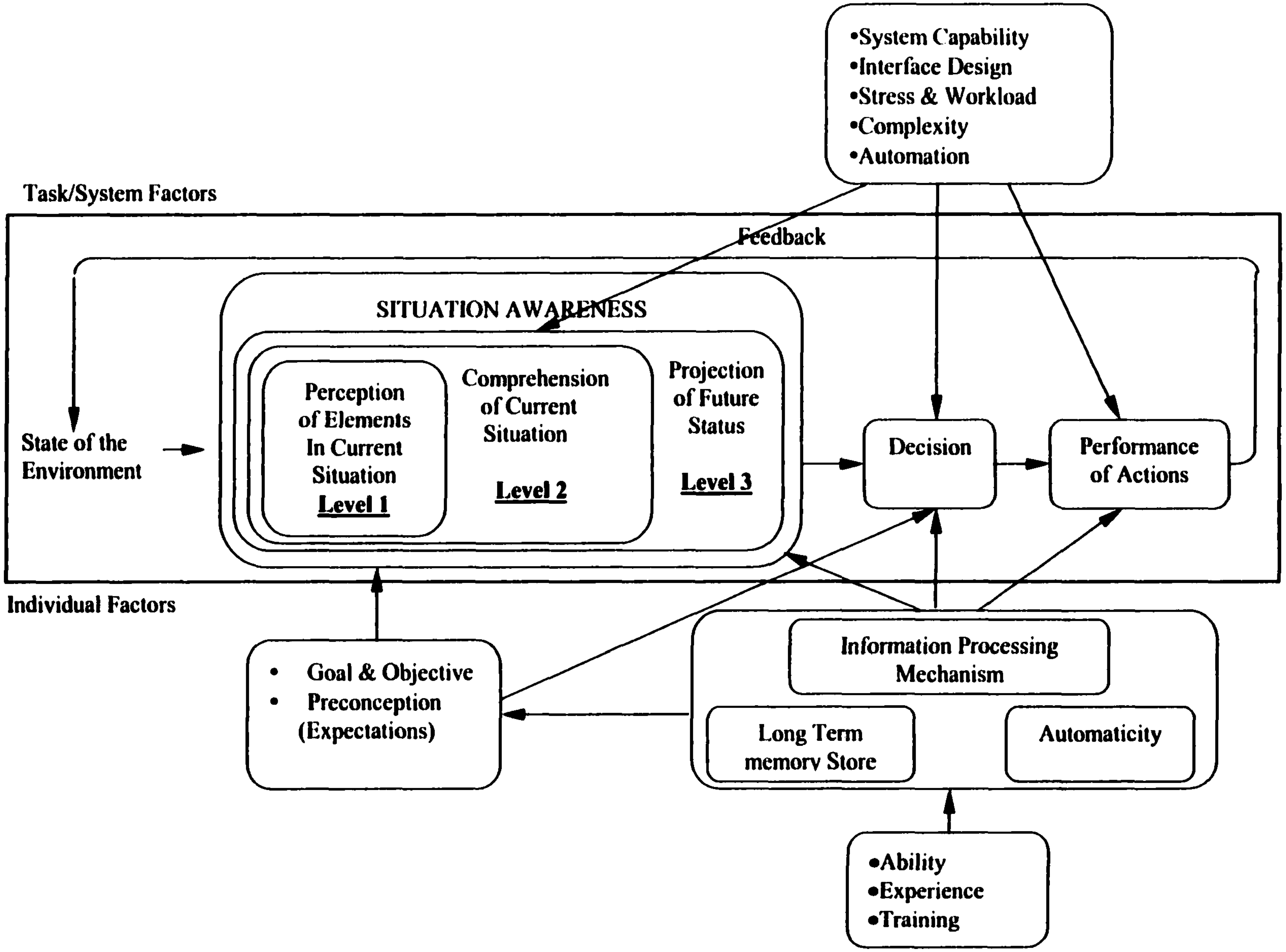

Figure 1. SA model (Endsley, 1995a)

Situation Awareness Model

Endsley (1995a) further categorized situation awareness into three different levels. The first level, called level 1 SA, is perception of the elements in the system. The next level, called level 2 SA, is the integration the pieces of perceived data and understanding their relevance to the goals. The highest level, called level 3 SA, is being able to forecast the future based on the elements just perceived and understood. The forecast provides the basis for making decisions.

Decision-making, and performing necessary actions are the steps occurring after situation awareness. Figure 1 shows the model of situation awareness proposed by Endsley (1995a).

The Importance of Situation Awareness

Situation awareness is needed for operators to perform tasks effectively (Endsley, in Endsley and Garland, 2000). To accomplish a goal, one must first, form situation awareness, then make decisions, and finally perform the necessary actions. These three steps occur continuously and are highly interrelated. The quality of every decision made is likely to be based on the completeness of the operators' situation awareness. Likewise, every action performed is based on the decision made (Endsely,1995a; Endsley, 1988). Therefore, task performance is based on the level of completeness of the situation awareness. In other words, it should be expected that poorer performance would result when situation awareness is less complete or less accurate.

The importance of situation awareness becomes greater as system complexity and dynamics increase. Elements in complex and dynamic systems vary across time, possibly at different rates, and are interdependent. In addition, current situation awareness affects the way the new information is perceived and interpreted (Sandom, 1999b; Endsley, 1988). Incomplete or inaccurate current situation awareness will lead to poorer situation awareness at later time. Operators therefore must continuously maintain high situation awareness. In the case of critical circumstances, where the operator must correctly react within only a limited amount of time, incomplete or inaccurate situation awareness can result in serious errors in decision making with disastrous consequences.

Situation Awareness Measurement

Endsley (in O'Brien, and Charlton, 1996) proposed a "process model", describing three stages required to reach a goal/sub-goal. The process begins with information assessment, where the operator gathers information considered to be critical in order to reach the goal. After the information is gathered, situation awareness is formed. The next step is making a decision and performing necessary actions to reach the goal. Methods used to measure SA vary depending on when the SA measurement is performed. According to the steps in the process model, SA measurements can be categorized into three approaches; Process indices, Direct Indices, and Behavior & Performance Indices.

"Process indices" are the measurements performed at the first stage. The idea behind this approach is that individuals will act differently to acquire SA, so the way they attend to the information, or process the information to develop SA should be different as well. Such differences therefore are used to indicate the operator's SA. The measurements can be taken in form of eye movements, information acquisition, or communication/verbalization. These methods reveal what information is viewed, or processed. However, the state of knowledge such as the amount of information in memory, the comprehension of the elements perceived, or whether the information is perceived correctly cannot be measured. Therefore, these techniques do not provide direct measurement of SA.

“Direct Indices” are the measurements performed at the second stage, when situation awareness is formed. The direct measurements are further divided into subjective and objective (questionnaire) techniques. For subjective techniques, the operator’s SA is rated by experienced observers, or by self-rating. Several subjective scales have been developed, for example, SART (Situation Awareness Rating Technique), and SARS (Situation Awareness Rating Scale). Alternately, objective techniques measure SA by questioning the operator to collect detailed information about their own perception of the current system status. The operator will be questioned either while performing the trial (on-line technique), at stops during the trail (SAGAT technique), or after the trial (posttest technique). The data collected can then be compared to what actually happened.

Behavior & Performance Indices are measurements performed at the last stage. The idea behind these methods is similar to Process Indices, that the operators will act differently depending on their SA. Therefore, the operator behavior and or performance such as communication frequencies, the amount of time needed to detect a faulty item, number of correct detections, etc., are used to indirectly indicate the operator’s SA.

Situation Awareness Global Assessment Technique (SAGAT)

SAGAT is a direct SA measurement, using a questionnaire technique to collect detailed information about the operator’s perception of the current system status. It was developed by Endsley (1988) to overcome the deficiencies caused by the other objective measurements, such as posttest or on-line techniques. Posttest techniques have operators complete detailed questionnaires after the simulated trial is complete. Operators are likely to forget some details of the events, and thus this method may capture SA only at the very end of the trial. On-line technique asks questions while operators are carrying out the trail in order to prevent the subjects from forgetting details of the events. However, asking in that manner may give operators clues of what to attend to next, or high workload tasks could prevent the operator from being able to answer detailed questions.

Unlike posttest and On-line techniques, SAGAT will pause the trial at randomly selected times, and quickly ask the operator for his/her perception of the situation at that point of time. After operators finish answering a set questions at each freeze, the trial continues. Stopping the trial and asking the operator questions could eliminate the problem of forgetting and intrusiveness. In addition, asking questions randomly will reduce the chance of operators' attention being bias toward any specific items. The questions cover almost every aspect of the situation including level 1, 2, and 3 components, and considers system functioning and status and relevant features of the external environment, thus it is a global measurement.

SAGAT has been proven valid in its predictive ability. Using the SAGAT method, Endsley (1990) found that the pilots who had the awareness of the presence of the opponent's aircraft were almost twice as likely to later kill that opponent's plane than the pilots who did not. Previous studies also showed that SAGAT is not intrusive. Endsley (1995b) showed that the number of stops during a trial had no significant affect on pilot performance. In addition, Snow and Reising (2000), used SA-SWORD, and SAGAT to measure pilots' situation awareness, and found that their flying performance was not affected by the method used.

Endsley (1998) stated that SAGAT is a sensitive measurement as it could detect the differences in pilot's situation awareness when forms of displays changed.

Factors Affecting SA

There are two groups of factors that affect an individual's SA: task/system factors, and individual factors. The availability of information needed and its form is one of the essential task factors that affect situation awareness. Other task/system factors such as workload, stress, complexity, and automation can greatly affect situation awareness. Individual factors include an ability of an individual to readily gather and interpret the information from the environment and form SA. This ability is determined by an individual's attention, and memory capacities. In addition, other factors such as experience, training, and expectations can also affect SA.

Task/System Factors

Complexity. System complexity refers to the degree of interconnection and interdependence among system components (Karwowski and Marras, 1999). As the complexity of the system increases, the possibility of having several side effects is higher due to the interdependence of system elements. In such a case, the possibility that the operator will become unaware of the actual state of the system could increase. In the other words, the more complex a system, the more difficult it is to maintain an adequate situation awareness.

Automation. As a result of automation, operator tasks have shifted from manual interaction with machines to remote operation via control system. Physical work is no longer needed, but considerable cognitive effort is needed to gather information related to system operation. Operators have more flexibility to access several levels of information. However, dealing with large amounts of information puts a high cognitive demand on the operators, and so they may easily lose awareness of the current system status. This suggests that automation can negatively affect operators' SA (Endsley, 2000; Sarter and Woods, 1995; Adams, Tenny, and Pew, 1995; Salvendy, 1997; Sandom, 1999a).

Workload. Workload is the stress, or demand, placed upon a system component. It may be cognitive or physical. (O'Brien and Charlton, 1996). "Workload is a multidimensional concept composed of behavioral, performance, physiological, and subjective components (Hart 1987), resulting from interaction between a specific individual and the demands imposed by a particular task" (Selcon, Taylor, and Koritsas 1991).

Workload is one of several factors hypothesized to affect the operators' situation awareness. Workload fluctuates over the operational period, and will hurt operators' situation awareness or task performance especially when it exceeds certain limits (O'Brien and Charlton, 1996; Endsley and Garland, 2000). Low workload coupled with boredom, fatigue, or sleep loss can have negative implications for human performance (Wickens and Hollands, 2000).

Workload is derived from several task-related factors. Limited time available could cause workload to increase as needed to prioritize tasks. The tasks considered to be of higher priority will be performed first, while the lower priority tasks may be ignored (Raby and Wickens, 1994). However, when evaluating task importance, operators could mistakenly neglect a high priority task, which could have dangerous consequences. Examples of such accidents include a failure of altitude monitoring in the case of the Eastern Airlines crash of the Lockheed L1011 in the Everglades (Wiener, 1977 in Raby and Wickens, 1994), or improper checklist procedure that caused the Northwest Airlines flight outside Detroit to crash (Wiener, 1989 in Raby and Wickens, 1994).

Information complexity can also increase operator workload, resulting in a less satisfactory performance (Svensson, Angelborg-Thanderz, Sjoberg, and Olsson, 1997), even if the information is theoretically not important for the task (Higgins and Chignell, 1988). When presented with unintelligible information, operators may need to expend considerable effort in interpreting and organizing before they can attempt to prioritize tasks. Morphew and Wickens (1998) found that the display which provided the most predictive information boosted flight safety, while reducing the pilot workload.

Feedback reflecting unsatisfactory performance could also cause operator workload to increase (Becker, Warm, and Dember, 1991). Operators having received negative feedback regarding their performance may try to work harder and more accurately to regain and maintain an acceptable performance level. This additional effort could result in an even greater workload for the operator (Borrensen, Bateman, and Malzahn,1988).

Individual Factors

Among the factors hypothesized to affect situation awareness, attention and memory are considered to be the major cognitive mechanisms important to develop SA (Endsley, 1988, 1995a; Endsley and Garland, 2000). Attention is needed to gather the information needed to form SA as operators use facts regarding the system to build mental models of the system in long-term memory. All three levels of situation awareness are supported by working memory; perceiving, comprehending, and projecting the future (Adams et al, 1995).

Attention and memory limitations. Below is the description of each of the attention and memory limitations.

(1) Knowledge of where targets will occur could reduce response time. Like a beam of light, there also exists a beam of attention, often described by the Spot Light Metaphor. When looking at a display, for example, the beam of attention can be moved from one area to another up to 24 degrees from the center point without moving the eyes (Posner, Snyder, and Davidson, 1980).

Posner, Nissen, and Ogden (1978) conducted an experiment to demonstrate this phenomenon. The experiment began with the subject staring at a fixed point. The subject then was cued with the likely spatial location of the target; either to the left or to the right of the fixed point. After the cue was presented, the target was presented. The subject responded to the presence of the target as quickly as possible. Only the trials in which the eyes stayed at the fixed point were included in the data analysis. The results showed that responses were faster when the target appeared corresponding to the location cue received, and slower when the target appeared in an unexpected location. The slower responses illustrated the need to shift the attention from one location to another to identify a target.

(2) Attention resource is limited. An individual's attentional resource is restricted, processing a limited amount of information at a time. A new item will be processed only if the pervious information has finished being processed. Raymond, Shapiro, and Arnell (1992) demonstrated this phenomenon. In their study, the subjects were shown a series of ten black letters and one white letter. The letters were presented as a Rapid Serial Visual Presentation (RSVP) with a rate of 11 items/second. The task was to identify the white target letter and the three letters presented immediately after the white letter. The results showed that the white target letter and the last letter in the stream were identified with the highest

probability, while the letters presented following the target were correctly reported with a significantly lower probability.

As the white target letter is being processed, the attention resource is being used and could not process the very next letter. Therefore, the chance that black letters would be identified would be greater if larger time gaps exist between the white letter and the appearance of the next black letters.

(3) Visual search time depends on targets. Visual search can be divided into two different types according to the relationship between the targets and the background. The first type of visual search is called a feature search, in which the targets are completely different form the background and, therefore, will always "pop-out" of the display. As a result, the subject can report the presence/absence of the target almost immediately. The second type is called a conjunctive search, in which the targets share at least one characteristic (i.e. color, or shape) with the background. The conjunct characteristics make the targets blend in with the distracters and increase the difficulty of identification. Therefore, it takes more time to report the presence/absence of a conjunctive target since they must be located before being able to confirm its presence/absence. That is, conjunctive search requires selective attention.

Treisma and Gelade (1980) conducted an experiment to demonstrate this paradigm. In the experiment, the subjects were asked to perform two task conditions. The first condition was to identify if a letter T_{blue} or X_{blue} was embedded among the array of T_{brown} and X_{green}. The second condition was to identify the letter T_{green} among the array of T_{brown} and X_{green}. In the first condition, the blue color makes the letter T distinguishable from the background letters. Conversely, in the second condition, the green color and shape of target letter T (T_{green}) may make it blend together with the background letters T_{brown} and X_{green}. The results showed that it took longer to identify the letter T_{green} when it was in the array of T_{brown} and X_{green}.

(4) Automatic responses interfere with information processing. Behaviors can become automated after being performed repeatedly a number of times, for example typing, driving, reading, playing a musical instrument, or bicycling. Once a behavior has been learned it becomes automatic to the point that it requires no or very little attention to repeatedly execute, and is likely to be difficult to stop from executing in the learned manner (Anderson, 1995). Automatic responses conflicting with the desired behavior, therefore, could interfere with other processing information. An experiment conducted by Dunbar and MacCleod (1984) demonstrated this phenomena.

In Dunbar and MacCleod 's experiment, the subjects were presented with a series of 100 words. The examples of words included in the study are "RED", "BLUE", and GREEN". The words could be printed in any color. The subjects were required to read each word, and then to identify the print color of the word the regardless of the word name (i.e. its meaning). The results showed that the subjects were much slower in identifying the print color when the print color was different from the word name. In addition, when the font color was deferent from the word name, subjects made more errors by identifying the word name as the font color. Because reading is an automatic process, the words read interfered with the subject's ability to identify the actual ink color.

(5) Location irrelevancy between stimuli and response slows down reaction time. The information-processing rate may be affected by the spatial relevancy between the stimuli and the response, resulting in a slow reaction time. For example, people will respond faster

and more accurately to stimuli that occur in the same relative location as the response (Coglab, 2000). Simon (1969) conducted an experiment to illustrate this characteristic. In his study, subjects were asked to respond to the presence of a 1000-cps monatural tone. A random sequence of tones was presented via the headphones the subject was wearing. Fifty percent of the tones were sent to the right ear and fifty percent to the left ear. The subject moved a control to either the left or the right of center depending upon which ear he/she heard the tone. There were two types of responses that subjects were required to make; moving the control toward the ear he/she heard the tone, and moving the control away from the ear hearing the tone. The results showed that reaction times when moving the control toward the source of the tone are significantly faster than when participants moved the control away from the source. Moving the control away from the source of the tone created a mismatch in direction of the stimuli and the response, so it slowed down the reaction time.

(6) Memory has a limited capacity. There is a limited amount of information that can he held in the short-term memory storage, a capacity called "memory span". In the other words, "memory span" is the number of items one can repeat back immediately after seeing a list of items (Anderson ,1995).

Ellis, Badderley and Miles (Baddeley, 1986) created several lists of digits, short words, long words, similar words and dissimilar words. These lists were presented to subjects one item at a time, with subjects reporting what they had just seen in the list. The greatest number of items remembered determined the memory span. The result showed that the digit span had the highest recall rate with the average of 6.7 items, while the similar word span had the lowest recall rate with the average of 3.42 items.

(7) False memory by mistaken recognition. There is evidence that humans may remember items that in fact have never happened, or remember it differently from the way it actually happened. This phenomenon is called false memory. Roedigler and MacDermott (1995) replicated a study done by Deese (1959) to illustrate the false memory paradigm. Subjects were shown a list with six items, and then recalled the items that were presented. Results showed that the probably of recalling a word not presented in the list was as high as 40%, with about 65% of words actually in the list recalled. Reporting words that were not in the list confirmed the existence of false memory.

(8) Ability to recall items presented in a list is limited by the location of the item in the list. Limitation of memory results in the ability to recall items more accurately from certain locations in the list. When subjects recall items from a list, in any order, the position of the item in the list has significant effect on the possibility that an item will be recalled. Roedigler and MacDermott (1995) showed that subjects could recall words placed at the beginning and at the end of the list more frequently than items in the middle

(9) The ability to identify one dimensional items is limited. When items with one dimension (i.e. weight, tones, or length) are evenly spaced, humans tend to be able to identify the first and the last item more reliably than items in the middle. Although research has confirmed the existence of absolute identification, a clear explanation of the cause of absolute identification has not been achieved (Shiffrin & Nosofsky, 1994).

Murdock (1960) conducted an experiment in which nine (numbered 1-9), evenly spaced tones were presented to the subjects once in ascending and once in descending order of magnitude. After the nine-tone presentation, one of the nine tones was randomly selected to be presented to the subject. The subject then identified the tone number he/she just heard.

The results showed that the subject could correctly identify the first tone and the last tone more readily than the tones in the middle.

Other individual factors. In addition to attention and memory, other individual factors also have been found to affect operators' situation awareness and performance. Huey and Boehm-Davis (1992) found that differences in gender and level of education affected the performance of operators of a simulated milk pasteurization plant. Torenvliet, Jamieson, and Vicente (1998) found that the interaction between a holist cognitive learning style and an interface based on the principle of ecological design was the strongest and most consistent predictor of performance. Gopher (1992) found that computer game experience can improve SA. Prior task experience or training were also found to improve SA and performance.

Summary

Situation awareness (SA) is the knowledge of current system elements status at an instant in time. It is an important concept for complex system operators as it is linked closely to task performance. The Situation Awareness Global Assessment Technique (SAGAT) is an objective measurement that has been proven valid in its predictive ability. Factors affecting situation awareness are system factors (i.e. automation, workload, and interface) and individual factors (i.e. attention, memory, and experience).

References

Anderson, J.R. (1995). *Cognitive Psychology and its Implications, 4thEd.* W.H. Freeman and Company NewYork.

Adams, M.J., Tenny, Y.J., Pew, R.W. (1995). Situation Awareness and the Cognitive Management of Complex Systems. *Human Factors,* **37***(1),* 85-104.

Becker, A.B., Warm, J.S., and Dember, W.N. (1991). Effects of Feedback on Perceived Workload in Vigilance Performing. In *Proceedings of the Human Factors Society 35th Annual Meeting,* (pp. 1491-1494). Santa Monica, CA: Human Factors and Ergonomics Society.

Borrensen, K.J., Bateman, R., and Malzahn, D. (1988). Techniques of Subjective Assessment: A comparison of the SWAT and Modified Cooper-Harper Scales. In *Proceedings of the Human Factors Society 32nd Annual Meeting,* (pp. 155-159). Santa Monica, CA: Human Factors and Ergonomics Society.

Coglab (2000). Simon Effect, wysiwyg://23/http://coglab.psyc...du/coglab/Labs/SimonEffect.html, Date accessed: 12/10/2000.

Deese, J. (1959). On the Prediction of Occurrence of Particular Verbal Intrusions in Immediate Recall. *Journal of Experimental Psychology,* **58**, 17-22.

Endsley, M.R., and Garland, D.J. (2000). Situation Awareness Analysis and Measurement, *LawrenceErlbaum Associates,* New Jersey.

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

Endsley, M.R. (1995a). Toward the Theory of Situation Awareness in Dynamic Systems, *Human Factors,* **37***(1),* 32-64.

Endsley. M.R. (1995b). Measurement of Situation Awareness in Dynamic Systems. *Human factors,* **37***(1),* 65-84.

Endsley, M.R., Situation Awareness, Automation & Free Flight, *http://vega3.uneec.eurocontrol.fr/endsley.htm,* Date accessed: 11/29/2000

Gopher, D. (1992). Development of Skill Trainers Based on Computer Games: Prospect and Issues. In *Proceedings of the Human Factors Society 36th Annual Meeting,* (pp. 1284-1290). Santa Monica, CA: Human Factors and Ergonomics Society.

Gugerty, L. 1998, Evidence from a Partial Report Task for a Forgetting in Dynamic Spatial Memory, *Human Factors,* **40***(3),* 498-508.

Higgins, T.J., Chignell, M.H. (1988). Mental Models: A Fifth Paradigm? In *Proceedings of the Human Factors Society 32nd Annual Meeting*, (pp. 145-149). Santa Monica, CA: Human Factors and Ergonomics Society.

Huey, B.M., and Boehm-Davis, D.A. (1992). Individual differences in the performance of a process control task, In *Proceedings of the Human Factors Society 36th Annual Meeting*, (pp. 897-901). Santa Monica, CA: Human Factors and Ergonomics Society.

MacCleod, C.M. (1992). The Stroop Task: The "Gold Standard" of Attention Measures. *Journal of Experimental Psychology: General*, **121**, 12-14.

Karwowski, W., & Marras, W.S. (1999). *The Occupational Ergonomics Handbook*, CRC Press, Washington, D.C.

MacCleod, C.M., and Dunbar, K. (1988). Training and Stroop-Like Interferences: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14**, 126-135.

Morphew, M.E., and Wickens, C.D. (1998). Pilot Performance and Workload Using Traffic Displays to Support Free Flight. In *Proceedings of the Human Factors Society 42nd Annual Meeting*, (pp. 52-56). Santa Monica, CA: Human Factors and Ergonomics Society.

Murdock, B.B. (1960). The Distinctiveness of Stimuli. *Psychological Review*, **67**, 16-31.

O'Brien, T.G., and Charlton, S.G. (1996). *Handbook of Human Factors Testing and Evaluation*, Lawrence Erlbaum Associates. New Jersey.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, **32**, 3-25.

Posner, M.I., Snyder, C. R.R., and Davidson, B.J. (1980). Attention and the Detection of Signals. *Journal of Experimental Psychology: General*, **109**, 160-174.

Raby, M & Wickens, C.D. 1994, Strategic Workload Management and Decision Biases in Avaiation, *The International Journal of Aviation Psychology* **4***(3)*, 211-240.

Raymond, J.E., Sharpiro, K.L., and Aemell, K.M. (1992). Temporary Suspension of Visual Processing in a RSVP Task : An Attentional Blink?. *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 849-860.

Roediger, H. L. III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 803-814.

Salvendy, G. (1997). *Handbook of Human Factors and Ergonomics* 2nd Edition. John Wiley&Sons. New York.

Sandom, C. (1999a). Operator Situation Awareness and System Safety, The Institution of Electrical Engineers, London.

Sandom, C. (1999b). Situation Awareness Through the Interface: Evaluating System in Safety Critical Control Systems. *IEE Conference Publication*, **463**, 119-124.

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.

Selcon, S.J., Taylor, R.M., & Koritsas, E. (1991). Workload or Situation Awareness?: TLX vs SART for Aerospace Systems Design Evaluation, In *Proceedings of the Human Factors Society 35th Annual Meeting,* (pp. 62-66). Santa Monica, CA: Human Factors and Ergonomics Society.

Shiffrin, R. M., & Nosofksy, R. M. (1994). 7 plus or minus 2: A commentary on capacity limitations. *Psychological Review*, **101**, 357-361.

Simon, J. R. (1969). Reaction toward the source of stimulation. *Journal of Experimental Psychology*, **81**, 174-176.

Stroop, J. R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology*, **18**, 643-662.

Svensson, E., Angelborg-Thanderz, M., Sjoberg, L., and Olsson, S. (1997). Information Complexity—Mental Workload and Peformance in Combat Aircraft. *Ergonomics*, **40***(3)*, 362-380.

Torenvliet, G.L.; Jamieson, G.A.; and Vicente, K.J. (1998). Making the Most of Ecological Interface Design: The Role of Cognitive Style. *Human Interaction with Complex Systems, Proceedings.*, Fourth Annual Symposium, 214 -225.

Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97-136.

Vidulich, M.A. (1992). Measuring Situation Awaress, In *Proceedings of the Human Factors Society 36th Annual Meeting,* (pp. 40-41). Santa Monica, CA: Human Factors and Ergonomics Society.

Wickens, C.D., and Hollands, J.G. (2000). Engineering Psychology and Human Performance 3rd Editiion, *Prentice Hall,* New Jersey.

Chapter 3. Attention and Memory Limitations of Engineering Students

A paper to be submitted to *Proceedings of The International Ergonomics Association*

Jaruwan Klamklay and Patrick E Patterson

Abstract

This study examined the attention and memory characteristics of engineering students. Eighty undergraduate engineering students performed nine psychological tests to measure their attention and memory limitations. The results showed that most of the attention and memory limitations of engineering students were similar to the limitations found in the previous studies. When comparing the results from this study with previous findings, it appears that the performance of engineering students was less impacted by the attention and memory limitations. Educational backgrounds and selected profession might affect how well individuals can overcome their attention and memory limitations during interactions with complex systems.

Introduction

With the advent of technological integration in industrial settings, modern complex systems are controlled via system control rooms having numerous control panels. Operators monitor system functions, diagnose malfunctions, and perform necessary corrective actions to bring a system back to normal values to prevent serious disruptions in production or hazardous situations. Such systems include air traffic control towers, airplane cockpits, power-generating plants, and manufacturing plants.

Due to the interconnection and interdependence among components within each complex dynamic system, a single component error or failure could result in numerous side effects. Paying attention to critical information, then making optimal decisions and selecting the right response, and finally performing correct actions are all crucial elements to achieving a desired performance. More importantly, these elements must occur in a very precise and timely fashion. Therefore, faster and more accurate information processing is critical to achieve a high level of performance and, more importantly, increased safety.

To achieve an optimal performance level, the displays and controls must enable fast and accurate information processing with system operators that are capable of efficiently and effectively operating the system. It is essential therefore to develop knowledge regarding operator information processing limitations to ensure that the system can be run at an optimal level. This knowledge could be applied to improve system design, to build systematic training programs, and to recruit operators for complex tasks.

Studies have shown that limitations of attention and memory are crucial elements governing human information processing, yet no previous studies have focused exclusively on the characteristics of specific subgroups, and then attempted to apply such knowledge to the system design, training programs, or operator selection processes. Previous studies have not questioned whether these limitations exist and whether a homogenous effect exists across individuals. Would it be possible that engineers could better overcome such limitations in complex systems than can non-engineers? The objective of this study was to determine the limitations of attention and memory of engineering students, and to determine if the performance of engineering students is less affected by such limitations. Below is the description of each of the attention and memory limitations.

Knowledge of where targets will occur can reduce response time. Posner, Nissen, and Ogden (1978 in Posner, Snyder, and Davidson, 1980) found that responses were faster when a target appeared according to the location cue received, and slower when the target appeared in an unexpected location. The slower responses were explained by the need to move the attention from one location to another to identify a target.

Attention resource is limited. An individual attentional resource is limited, processing a limited amount of information at a time. A new item will be processed only if the pervious information has finished being processed (Raymond, Shapiro, and Arnell, 1992). As a target is being processed, the attention resource is being used and may not be able to process the next information. Therefore, the chance that non-target information would be identified is greater if larger time gaps exist between the target and the appearance of the subsequent non-target information.

Visual search time depends on the target. Visual search can be divided into two different types according to the relationship between the targets and the background. The first type of visual search, feature search, in which the targets are completely different form the background and, therefore, will always "pop-out" of the display. As a result, the subject can report the presence/absence of the target almost immediately. The second type is conjunctive search, in which the targets share at least one characteristic (i.e. color, or shape) with the background. The conjunct characteristics make the targets blend in with the distracters and increase the difficulty of identification. Therefore, it takes more time to report the presence/absence of a conjunctive target since they must be located before being able to confirm its presence/absence. In the other words, conjunctive search requires selective attention (Treisma and Gelade, 1980).

Automatic responses interfere with information processing. Behaviors can become automated after being performed repeatedly a number of times, for example typing, driving, reading, playing a musical instrument, or bicycling. Once a behavior has been learned it becomes automatic to the point that it requires no or very little attention to repeatedly execute, and is likely to be difficult to stop from executing in the learned manner (Anderson, 1995). Automatic responses conflicting with the desired behavior therefore could interfere with other processing information (Dunbar and MacCleod, 1984).

Location irrelevancy between stimuli and response slows down reaction time. The information-processing rate may be affected by the spatial relevancy between the stimuli and the response, resulting in a slow reaction time. For example, people will respond faster and more accurately to stimuli that occur in the same relative location as the response (Simon, 1969).

Memory has a limited capacity. There is a limited amount of information that can he held in the short-term memory storage, a capacity called "memory span". In the other words, "memory span" is the number of items one can repeat back immediately after seeing a list of several items (Anderson ,1995; Ellis, Badderley and Miles in Baddeley, 1986).

False memory created by mistaken recognition. There is evidence that humans may remember items that in fact never happened, or remember them differently from the way they actually occurred, a phenomenon termed false memory (Deese, 1959; Roedigler and MacDermott;1995).

Ability to recall items presented in a list is limited by the location of the item in the list. Limitation of memory results in the ability to recall items more accurately from certain locations in the list. When subjects recall items from a list, in any order, the position of the item in the list has significant effect on the possibility that an item will be recalled (Roedigler and MacDermott , 1995).

The ability to identify one dimensional items is limited. When items with one dimension (i.e. weight, tones, or length) are evenly spaced, humans tend to be able to identify the first and the last item more reliably than items in the middle. Although research has confirmed the existence of absolute identification, a clear explanation of the cause of absolute identification has not been achieved (Murdock, 1960; Shiffrin and Nosofsky, 1994).

The objective of this study was to examine the information processing characteristics of engineering students, and to determine if engineering students are better in overcoming such limitations. Nine psychological tests duplicated from previous studies were used to quantify the limitations of their attention and memory. These tests include Attentional Blink, Spatial Cuing, Visual Search, Stroop Effect, Simon Effect, Absolute Identification, False Memory, Memory Span, and Serial Position. The first hypothesis was that the effects of attention and memory limitations exist among engineering students. With this in mind, the following results were expected for the nine aforementioned tests.

Attentional Blink: The detection of the second target should increase as the separation between the first and the second target increases.

Spatial Cueing: Compared to the non-cue case, the subject should respond faster when the target appears in the cued location, and slower when the target appears in a non-cued location.

Visual Search: Conjunctive search times increase as the number of distracters increase, especially in the case of conjunctive "absent" searches. However, the number of distracters should have little effect on search time for the feature condition.

Stroop Effect: The reaction time is longer when the word names and colors are different.

Simon Effect: The reaction time should be faster, and fewer errors would be made when the location of the stimuli matched the location of the response.

Serial Position: The percentage of correctly recalling the first and last few items in the list will be higher than of the items in the middle.

Absolute Identification: The ability to identify the one dimensional items in the list should be better for items closer to the ends of the list.

Memory Span: The memory span should be 7 ± 2 items.

False Memory: The frequency of selecting the special distracters should be larger than the frequency of selecting items that were not presented in the list but smaller than the frequency of selecting the items that were in the list.

The second hypothesis was that, by comparing the results from the current study with results from similar previous studies or theories, attention and memory have less effect on the performance of engineering students. Due to insufficient reference data, a direct statistical comparison could not be performed. The following results from the comparison were expected.

- *Attentional Blink*: Engineering students will have higher detection percentage of the second target.
- *Spatial Cueing*: Engineering students will have shorter reaction time amongst the three task conditions: valid cued, invalid cued, and non-cued.

- *Visual Search*: Number of distracters will have less impact on engineering students conjunctive search times. In addition, distracters had no effect on search time for the feature condition.
- *Stroop Effect*: Engineering students will have a smaller increase of reaction time caused by the difference between the word names and colors.
- *Simon Effect*: Engineering students will have a smaller increase in reaction time and proportion error caused by the irrelevancy between the location of the stimuli and the response.
- *Serial Position*: Engineering students will have a higher probability of correctly recalling items in the list.
- *Absolute Identification*: Engineering students will have a higher probability of correctly identifying one-dimensional items in the list.
- *Memory Span:* Engineering students will have a larger memory span.

Methods

Participants

Eighty undergraduate engineering students (25 female and 55 males) ages between 19-25 years from various disciplines were enrolled on a voluntary basis to participate in this study. The study was evaluated and approved by the Iowa State University Human Subject Committee.

Apparatus

Nine classical psychology tests were administered to the participants. Attention Blink, Spatial Cuing, Visual Search, Stroop Effect, and Simon Effect tests were used to measure participant attention characteristics. The False Memory, Memory Span, Absolute Identification, and Serial Position tests were used to measure the participant memory characteristics. These tests were reproduced from previous studies and made available to public via the World Wide Web by the Psychology Department at Purdue University.

Procedures

Each participant performed the nine tests in random order, with a short break between tests. The followings are the brief procedures of each test.

Attentional Blink: The subject watched several series of ten letters, with a new letter overwriting the previous letter. Each series contained the first target letter, and/or the second target letter. The separation between the presence of the first target and the second target varied between series. The subject reported the presence of both target letters. The percentage of time the subject could correctly identify the first and second targets was recorded as a function of separation time between the targets.

Spatial Cuing: The operator was required to respond to the presence of targets as quickly as possible under three possible conditions: (1) when a cue correctly identified a location a target (2) when a cue incorrectly identified a location of a target, and (3) when there was no cue. The reaction times under these three conditions were recorded.

Visual Search: The subject performed two types of searches: Conjunctive search, and Feature search. For the conjunctive search, the subject determined whether a green circle was present among green squares and blue circles. For feature search, subjects determined whether a green circle was presented among the blue squares. The non-target items, such as green squares and blue circles, were distracters. Reaction time as a function of the number of distracters was recorded.

Stroop Effect: The subject was required to identify the font colors of the words "BLUE", 'RED", and "GREEN". The words could be written in the font color blue, red, or green. The reaction times to identify the font color when the font color and the word name were the same, and when they were different, were recorded.

Simon Effect: The subject was required to respond to the presence of the target as quickly possible under four different conditions. The four conditions were: (1) When the target was presented on the right, and the response key was located on the right, (2) When the target was presented on the left, and the response key was located on the left, (3) When the target was presented on the right, and the response key was located on the left, and (4) When the target was presented on the left, and the response key was located on the right. The reaction times and the proportion of errors from each condition were measured.

Memory Span: Several series, each containing only one of the following types of stimuli: numbers, letters sounding different, letters sounding the same, short words, and long words, were randomly presented to the subject. After completing each series, the subject recalled the items in the series. The numbers of items the subject could recall as a function of stimulus type were recorded.

Serial Position: The subject was presented with several series of ten letters. In each series, the subject recalled the letters that were in the list in any order. The average percentage of time a subject correctly recalled an item at each position in the sequence was recorded.

Absolute Identification: The subject was first presented with a series of seven lines, starting from the shortest line (line 1), to the longest line (line 7). After previewing the seven lines, one of the seven lines was randomly presented to the subject and then repeated several times. Each time the subject identified which line was presented. The average percentages of time a subject could correctly identify line 1 to line 7 were recorded.

False Memory: The subject was presented with a sequence of words. Upon sequence completion, the subject chose the words in the sequence just seen from a pool of words. The pool contains; (1) words that were in the sequence, (2) words that were not in the sequence, and (3) special words very similar to the words that were in the sequence. The percentage of time the subject chose each type of words were recorded.

Results

The results from each test are described below. The JMP statistical program (JMP 4.0.2) was used to perform the t-test analysis on each test. A .05 significant level was used throughout the entire analysis.

Attentional Blink: The results (Table 1) show that as the separation time between the first and the second targets increased, the percentage of time the operator could detect the second target also significantly increased ($p < .0001$). However, increasing the separation between targets from 200 milliseconds to 400 milliseconds did not improve the ability to detect the second target.

When compared to the results from the current study with Raymond et al (1992), it appears that engineering students might have slower information processing rate. At the 800 millisecond separation between the first and the second target, the engineering students participating in this study detected the second target letter with the possibility of 60%. However, in Raymond et al experiment, when separation was at 727 ms, the probability of reporting the successive target was as high as 80%.

Table 1. The average percentage of second target detection as a function of the separation time (milliseconds) between the first and second target

% second target detection	**Separation Time (milliseconds)**				
	0	**200**	**400**	**600**	**800**
Present study	6.19	38.38	38.00	53.00	59.81
Raymond et al's*	0**	17.60**	27.5**	39.60**	90.20**

** Participants were 3 male and 2 female university students and staff, ages between 22-39 years old*

***Values are interpolated*

Spatial Cuing: The results (Table 2) show that the presence of the cue did not affect the reaction times to the presence of the target, regardless whether the valid cue was a valid cue, an invalid cue, or no cue was presented ($p < .05$).

However, the results from Posner, Nissen, and Ogden (1978, found in Posner, 1980) showed that a valid cue resulted in an approximately 20 millisecond shorter reaction time, and the invalid cue resulted in an approximate 40 millisecond longer reaction time. This suggests that engineering students were less dependent on the knowledge of spatial cue location of the target in reorienting the attention to search for the target.

Table 2. The average reaction times (milliseconds) under three cue conditions

Reaction times (ms)	**Cue Conditions**		
	Valid cue	**Invalid cue**	**No cue (Neutral)**
Present study	351.41	377.98	356.58
Posner et al's*	240	300	260

**A total of 23 participants participated in this study, but no further details were provided*

Visual Search: The average conjunctive search time increased linearly with the number of distracters, with the slope averaging 14.7 ms for the target present condition and

24.9 ms for the target absent condition (Tables 3a and 3b). The feature search time was not affected by the number of distracter, with the slopes averaging .16 ms for the target present condition, and .01 ms for the target absent condition.

Table 3a. The average visual search times (milliseconds) under search conditions, and number of distracters.

Search Type	**Average visual search times (ms)**		
	Number of distracters		
	4	**16**	**64**
Conjunctive Present	714.2394	983.3971	1317.7100
Conjunctive Absent	886.6895	1316.289	1985.6720
Feature Present	563.5238	599.1645	606.8706
Feature Absent	663.9331	692.0690	692.9050

Table 3b. The slope of visual search times (milliseconds) on number of distracter

Search Types	**Slope**	
	Present study	**Treisman and Gelade's***
Conjunctive Present	14.7	28.7
Conjunctive Absent	24.9	67.1
Feature Present	.16	3.1
Feature Absent	.01	25.1

**Participants were 6 male and 2 female members of the Oxford subject panel ages between 24-29 years old*

It appears that number of distracters, or the presence of a target, has less effect on engineering students' visual search times. This is supported by the much larger average slopes found in Treisman and Gelade (1980), whose participants had a slope of 28.7 for the target present conjunctive search, 67.1 for the target absent conjunctive search, 3.1 ms for the target present feature search, and 25.1 ms for the target absent feature search.

Stroop Effect: The results (Table 4) show that the difference of font name and font color significantly increased the average reaction time, p = .0397. The automatic reading process had less effect on engineering students than average participants. The result from the current study showed that the difference between font color and word name resulted in a 7.3% time increase. However, the result from Stroop showed that such difference caused the reaction time to increase up to 74%.

Table 4. The average reaction times (milliseconds) to identify the font color

Cases	**Reaction times (ms)**	
	Present study	**Stroop's***
When font color and word name are the same	766.10	630.3
When font color and word name are different	821.97	1100.3

**Participants were 14 male and 56 female undergraduate college students*

Simon Effect: The location of the respond key relative to the location of the stimuli effected the reaction time and proportion error (Table 5). When the location of the target and the response key did not match, a longer reaction time resulted with a larger proportion error. Using different hands (i.e. left or right) to activate the response key did not create any differences in reaction times or proportion errors.

When our results are compared to those of Simon (1969), they suggest that cue irrelevancy may have has less impact on engineering students. In the current study, irrelevancy caused the reaction times to increase by approximately 9% for incongruent left, and 8.4% for incongruent right. In Simon's study, irrelevancy caused the reaction times to increase by approximately 11.6% for incongruent left, and 13.6% for incongruent right.

Table 5. The average reaction times (milliseconds) to respond to the target, and proportion error

Cases	**Reaction times (ms)**		**Proportion error**	
	Present study	**Simon's***	**Present study**	**Simon's***
Congruent Left	482.1216	480	.0360	-
Congruent Right	476.2549	479	.0290	-
Incongruent Left	525.6738	536	.0785	-
Incongruent Right	516.1032	544	.0725	-

**Participants were 32 male and 32 female undergraduate students, ages between 18-24 years old*

Memory Span: The results (Table 6) show that the memory spans of engineering students were approximately in the range of 7 ± 2. The subjects demonstrated the largest memory span on numbers and the shortest span on long words. The letters sounding different, the letters sounding the same, and the short words were in the middle rank memory span.

Table 6. The average numbers of items recalled

Stimuli	**Average number of item recalled**	
	Present study	**Badderly's***
Numbers	6.925	6.70
Letters sounding different	6.275	4.70
Letters sounding similar	5.400	4.28
Short words	5.700	3.42
Long words	4.500	3.65

**Participants were a group of boys*

When compared our results to those of Badderly (1986), it appears that the engineering students had larger memory span for almost all type of stimuli. The current study found average spans of 6.9, 6.3, 5.4, 5.7, and 4.5 for digits, letters sounding different, short words, letters sounding the same, and long words respectively, while Badderly found the spans of 6.7, 4.7, 4.28, 3.42, and 3.65 for the same set of stimuli.

Serial Position: The results (Table 7) show that the first three words in the list were correctly recalled more frequently than the words in the middle of the list. The last two words in the list were also recalled with a high probability. The engineering students participating in this study performed more poorly in recalling items at the beginning and at the end of the list. However, they performed better in recalling items in the middle of a list.

Table 7. The average percentages of item correctly recalled

% Correctly identified	**Position in list**											
	1	2	3	4	5	6	7	8	9	10	11	12
Present study	75.6	70.6	66.8	64.6	64.2	59.6	62.0	62.3	66.1	67.7	-	-
Roedigler and MacDermott's*	86	77	64	50	40	39	40	58	65	78	88	98

**Participants were 36 undergraduate students taking a Psychology class*

Absolute Identification: Subjects were more accurate in identifying the lines close to the extreme lengths than the lines in the middle of the range. For example, the ability to identify the lines 1 and 7 was greater than the ability to identify lines 2 and 6. The ability to identify lines 2 and 6 was greater than the ability to identify lines 3 and line 5 (Table 8a).

Murdock (1960) stated that the ability to identify stimuli depends on the stimulus distinctiveness. He defined stimulus distinctiveness (D) as the sum of the differences between it and all other stimuli in the group. D% is the summation of D_i divided by the total D, used to predict the chance the stimulus D will be detected. Table 8b shows the comparison of D% based on Murdock's theory, and the actual D% from our experiment. In comparison to Murdock's results, the engineering students participating in this study performed more poorly in identifying items at the beginning and at the end of the list. However, they performed better in identifying items in the middle of a list.

Table 8a. The average percentages of items correctly identified

	Line number						
	1	**2**	**3**	**4**	**5**	**6**	**7**
% Correction	70.25	61.25	58.38	53.75	56.75	62.25	68.63

Table 8b. %D based on Murdock's theory and the actual values of this study

% D	**Line Number**						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Present study	16.3	14.2	13.5	12.5	13.2	14.4	15.9
Murdock's	18.8	14.3	11.6	10.7	11.6	14.3	18.8

False Memory: The results (Table 9) show that the subjects chose the words that were actually in the sequence more than the words that were not. Subjects were also more likely to choose the distracter words than the normal words not in the sequence. Surprisingly, the percentage of time the subjects mistakenly chose the special distracter words was as high as the percentage of time they could correctly identify the words actually in the list.

Table 9. The average percentages of items selected

Stimuli	**Average % selection**	
	Present study	**Roediger and McDermott's***
Words in the list	81.4779	65.00
Normal words not in the list	2.8906	-
Special distracter	78.9115	40.00

**Participants were 36 undergraduate students taking a Psychology class*

The results from this False Memory experiment were quite different from those reported by Roediger and McDermott (1995). In their study, the probability of reporting the words presented in the list was 65%, and the probability of reporting the special distracter words was 40%. In this study, the probability of reporting the words presented in the list was 81.5%, and the probability of reporting the special distracter words was 79%.

Discussion and Conclusion

As expected, most of the attention and memory characteristics of engineering students were similar to the limitations found in the previous studies. Only the results from the Spatial cueing and the visual search tests were slightly different from previous findings. This study showed that previous knowledge of spatial location of the target did not affect the response time to the target, and that number of distracters did not affect the feature search time.

When comparing the results from this study with previous findings, although statistical tests could not be performed due to insufficient reference data, it appears that the performance of engineering students was less impacted by the attention and memory limitations. For example, engineering students could more quickly reorient their attention to locate the target regardless whether a spatial cue was provided or not. The effect of automatic responses to the meaning of the words, and the effect of location irrelevancy between the stimuli and the response were found to have less of an impact on engineering students. Engineering students were also found to have larger memory spans, especially in remembering digits. Engineering students were poorer in identifying items at the beginning and at the end of the list, yet they were better at identifying items in the middle of the list. Engineering students in this study however, had larger false memories.

The results found encourage further study to ascertain the differences of limitation differences of attention and memory across different groups of individuals. Complex system designers should carefully investigate the limitations of attention and memory of target users to their designs. This information could also be applied in developing systematic training courses, and operator selection methods.

In summary, this study collected information regarding the attention and memory limitations of undergraduate engineering students. The attention and memory limitations of engineering students who participated in this study were similar to that of non-engineer participants of previous studies, however engineering students seemed to be less impacted by such limitations. The collected data suggested that educational backgrounds and selected professions may provide some indication of how well individuals can overcome their natural attention and memory limitations. Future research should compare the results of this study with the results obtained from subjects with different educational backgrounds (i.e. non-engineering).

References

Anderson, J.R. (1995). *Cognitive Psychology and its Implications, 4thEd.* W.H. Freeman and Company NewYork.

Baddeley, Alan. (1986). *Working Memory.* Oxford University Press New York USA.

Deese, J. (1959). On the Prediction of Occurrence of Particular Verbal Intrusions in Immediate Recall. *Journal of Experimental Psychology*, **58**, 17-22.

Miyake, Akira., and Shah, Priti. (1999). *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control.* Cambridge University Press New York USA.

MacCleod, C.M. (1992). The Stroop Task: The "Gold Standard" of Attention Measures. *Journal of Experimental Psychology: General*, **121**, 12-14.

MacCleod, C.M., and Dunbar, K. (1988). Training and Stroop-Like Interferences: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14**, 126-135.

Murdock, B.B. (1960). The Distinctiveness of Stimuli. *Psychological Review*, **67**, 16-31.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, **32**, 3-25.

Posner, M.I., Snyder, C. R.R., and Davidson, B.J. (1980). Attention and the Detection of Signals. *Journal of Experimental Psychology: General*, **109**, 160-174.

Raymond, J.E., Sharpiro, K.L., and Aemell, K.M. (1992). Temporary Suspension of Visual Processing in a RSVP Task : An Attentional Blink?. *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 849-860.

Roediger, H. L. III, and McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 803-814.

Shiffrin, R. M., and Nosofksy, R. M. (1994). 7 plus or minus 2: A commentary on capacity limitations. *Psychological Review*, **101**, 357-361.

Simon, J. R. (1969) Reaction toward the source of stimulation. *Journal of Experimental Psychology*, **81**, 174-176.

Stroop, J. R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology*, **18**, 643-662.

Treisman, A. M., and Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12**, 97-136.

Wickens C.D., and Hollands, J.G. (2000). *Engineering Psychology and human Performance* 3rd Ed, Prentice-Hall Inc. New Jersey.

Chapter 4. The Effect of Automation, Pace, and Task Duration on Operator Situation Awareness and Performance

A paper to be submitted to *Human Factors*

Jaruwan Klamklay and Patrick E Patterson

Abstract

This study examined the effect of automation, pace, and task duration on situation awareness and task performance of operators performing a simulated complex task. A simplified pilot task required operators to monitor system status and manage the system fuel resource. Eighty participants performed the simulated task under different task paces, duration, and automation levels. Speed and accuracy were used to measure the task performance and the Situation Awareness Global Assessment Technique (SAGAT) used to quantify situation awareness. A multiple regression model was created for each of the response variables to determine the effect of automation, pace, and task duration. The results showed that automation reduced situation awareness, and yielded mixed effects on task performance. High task pace improved situation awareness yet had mixed impacts on task performance. Longer task duration had negative impacts on task performance, yet had no effect on situation awareness.

Introduction

The concepts of situation awareness (SA) and task performance are generally used in system design and evaluation. The system should contribute a satisfactory level of situation awareness as well as performance.

System automation can reduce workload when time stress is high, or when cognitive effort is high (Wickens and Holland, 2000 page 542). However, including automation in the system could also negatively impact operator situation awareness in three possible ways: trust, information processing, and feedback (Endsley, 2000). The over-trusting of system automation could reduce the vigilance performance, while under-trusting can result in over checking. In manual control, the information is active, while the automated system is passive. Therefore, the sophisticated automation algorithms of automated systems may confuse operators, inhibiting them from understanding of what the automated system is actually doing (Wickens and Holland, 2000). The loss of feedback, which always occurs in automated systems, could also harm higher level of situation awareness. Incomplete SA then may cause performance degradation.

Endsley and Kaber (1999) examined the effect of several automation levels on operator performance and situation awareness. Monitoring the displays, generating processing options, selecting an optimal option, and implementing the selected option were the functions allocated to the operator and/or the computer of the system. The results showed that allowing an operator to select the option, while allowing joint human/machine to monitor the displays, generate options, and implementing performance, could best improve task performance under normal system condition. However, less automation involvement allowed the system operator to more quickly recover the system from a failure. Poorer SA occurred when less automation was applied. However, if a system is fully automated and an operator is completely removed from the system operation, it may lead to a high possibility of out-of-the-loop performance decrements.

Time is considered to be another critical factor for SA, as SA is built up over time (Endsley, 1995). As technology advances, systems will be able to operate at a faster rate and in more complicated ways. When time constraints are presented, decision makers may fail to employ strategies that enable them to select the best alternative (Johnson et al 1993). Operators with the ability to acquire and maintain high situation awareness of the system therefore are critical. Raby and Wickens (1994) found that increasing flight pace and communication frequency resulted in more errors in most performance measures. Hancock and Caird (1993) found that increasing target shrink rate resulted in faster reaction time (RT) but increased movement time (MT). In addition to the effect of automation and fast pace, long task duration may cause fatigue and boredom, resulting in a loss of situation awareness and poor performance (Hankey and Dingus, 1990).

Studies have been conducted regarding the effect of task factors on situation awareness and performance. However, most of these studies focused on aviation systems, and used highly experience pilots as the subjects. Since SA and task performance are domain specific, each system may require its own method of measurement. Therefore, findings from previous studies may not be generalized across different systems and system operators.

The objective of this study was to determine the effect of automation, pace, and task duration on the novice operator of a simulated complex system. The results obtained could help clarify the SA construct, and can be applied to system design and evaluation.

Methods

Participants

Eighty undergraduate engineering students (25 female and 55 males) from various disciplines were enrolled on a voluntary basis. The study was evaluated and approved by the Iowa State University Human Subject Committee.

Task

This study used the Multi-variable Attribute Task battery for Human Operation Workload and Strategic Behavior Research (Comstock and Arnegard, 1992) to simulate a simple complex system. Each operator of the system was required to perform two sub-tasks. The first sub-task was system monitoring, in which the operator monitored two lights and four vertical scales for signs of system abnormalities. Under normal conditions, the left light would always be green. However, if a system abnormality occurred, the green light would go out. The operator responded to the disappearance of the green light by pressing the assigned key as soon as possible. The right light was normally off; however, if a system abnormality occurred a red light would turn on in that position. Similarly, once the operator noted the red light was on, he/she pressed the assigned key as soon as possible. The operator would see feedback as the green light would immediately turn back on, and the red light would immediately turn off.

In addition to the lights, the operator also monitored the vertical scales below the lights. Each scale had a yellow pointer that could fluctuate one unit below to one unit above the centerline. If a system fault occurred, the corresponding scale would shift out of the normal range. Regardless of the abnormal shifting direction, the operator could correct the fault by pressing the assigned key. The feedback to a correct response was given by the presence of a yellow bar at the bottom of the scale that indicated an out of range condition, and a return to the center of that scale pointer. If the operator did not notice the fault events (i.e. the lights, the scales), the situation would return to normal condition after a selected time-out period.

The second sub-task was resource management. This sub-system contained six fuel tanks connected by pipelines. The main tanks were tanks A and B, each having 4000-unit capacity. The supply tanks C and D contained a maximum of 2000 units each, and two additional unnamed supply tanks had unlimited capacities. The system used the fuel from tanks A and B, so fuel levels in these two tanks continuously decreased while the system was operating. The operator's task was to maintain the fuel in tanks A and B at their specific levels, as well as trying to keep tanks C and D full. Both numbers underneath tanks A, B, C, and D, and the green levels in the tanks, represented the current amount of fuel in the tanks. To meet the goal, the operator used the eight pumps provided to transfer fuel between tanks. The pumps transferred fuel from one tank to others in the directions indicated by arrows on the display. Keys numbered 1 to 8 were used as toggle switches to turn the corresponding pumps On/Off. No interactions from the operator were required if resource management was in its automatic mode.

The experimenter could set several system parameters including the disappearance of the green light, the appearance of the red light, the out-of-range fluctuation of the scales, time-out periods, resource management automatic mode, pump failures, service time to the failed pumps, pump flow rates, and the fuel consumption rates of tanks A and tank B.

Task performance was measured from the average reaction times to detect and correct system malfunctions (i.e. the scales, the green light, and the red light), proportions of misses (i.e. number of uncorrected malfunctions/ total number of malfunctions), and the average deviation of fuel levels of tanks A, B, C, and D from their target levels.

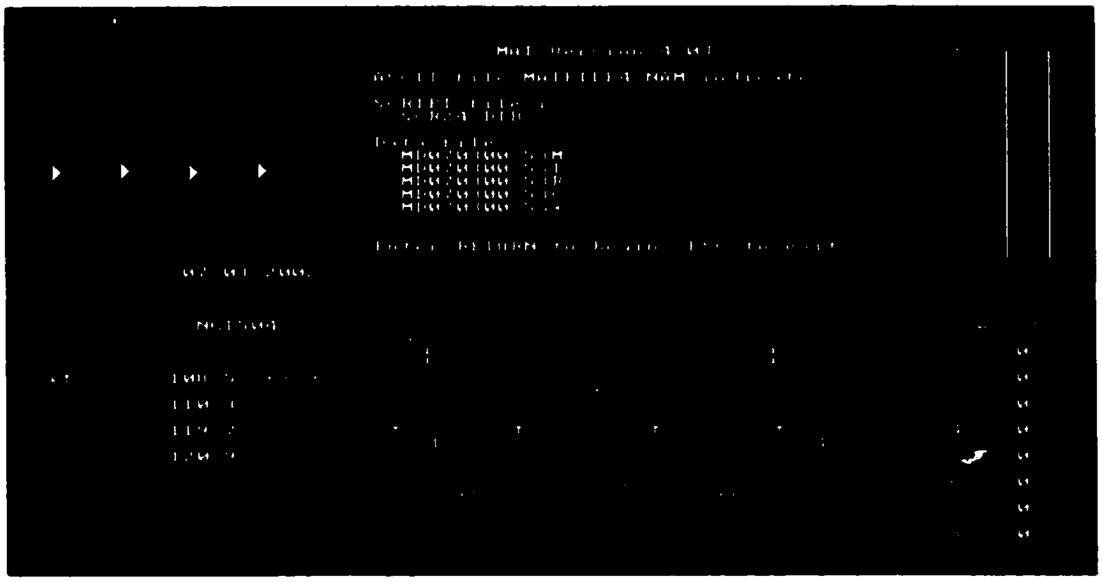

Figure 1. The MAT task

Each operator's SAGAT score was calculated by comparing the operator's answers to what actually happened. Each query was graded independently. Seven groups of questions presented to the subjects were: (1) the current status of the lights (on/off), (2) if the lights were currently in the desired conditions, (3) if the scales were currently in the desired conditions, (4) if the fuel amounts in the tanks were in the desired ranges, (5) the current status of the pumps (i.e. malfunctioning, operating, and not operating), (6) the current tasks of the pumps (i.e. supplying which tanks), and (7) future plans for the pumps (which pumps to turn on).

Questions were either Yes/No questions, or multiple answer questions. For Yes/No questions, if the question was answered correctly the score would be one, otherwise zero. For the questions that contain more than one possible answer, the score was determined by the number of correct answers divided by the summation of the number of correct and incorrect answers. Possible answers that were not chosen were considered incorrect answers. The scores for this type of question ranged between zero and one.

Task factors

Three task factors were included in this study: pace of work, task duration, and automation level.

Pace. Task parameters such as number of fault events, fault timeout periods, system fuel consumption rates, pump capacities, and pump reliability were used to differentiate the two pace levels of work. The first level of work pace contained fewer fault events with each having an extended timeout period. This longer timeout period allowed the operator more time to detect and correct the faults before the system resets the fault condition back to normal. At the slower pace, the system consumed less fuel, resulting in slower consumption of the fuel from both tanks A and B. In addition, the pumps had variable capacity levels enabling the operator to manage and maintain the amount of fuel at the desired levels. Also, there were fewer pump failures, and more time was required to repair the pump.

The faster pace contained a greater number of system faults, with each fault having a shorter timeout period, giving the operator less time to detect and correct the faults. The system consumed more fuel and resulting in the faster consumption rates of fuel levels in tanks A and B. These pumps had higher capacities, yet created more difficulties in managing them to reach and maintain the amount of fuel at the desired levels. The number of pump failures increased, with failed pumps requiring less time to be fixed.

Task Duration. Two levels of task duration (20-minute and 40-minute) were used in the simulation task.

Automation Level. Two levels of automation, 0% automation, and 50% automation, were included in the experimental trials. At 0% automation, both the system monitoring sub-task and the resource management sub-task were manually performed. For the higher level of automation, the operator manually performed the monitoring sub-task, while the resource management sub-task was automatically performed by the system. However, the operator was required to manually perform the resource management sub-task if the automated system failed.

Experimental design

The experiment used a full factorial design, with 2 levels of paces, 2 levels of task duration, and 2 level of automation, resulting in 8 possible treatments. Each subject performed two different simulation treatments. The dependent variables for this study were the operators' performance and situation awareness.

Procedure

The first session began with a 15-minute introduction and practice, allowing the participants to become familiar with the MAT program, the NASA-TLX workload rating scale, and the SAGAT measurement. MAT system components and their fundamental characteristics were explained. The subjects were informed in advance to expect system failure events such as the disappearance of the green light, the appearance of the red light, the deviations of the scales, pump failures, and automation mode failures. Each subject practiced responding to the event faults and controlling the pumps via the computer keyboard. Next, the NASA-TLX self-rating procedure was introduced including the meaning of each sub-scale. Each subject learned how to use the NASA-TLX scale, followed by a practice session. The SAGAT procedure was then introduced to the subjects. The subjects were informed to expect freezes and questions during the experimental session. The subjects were encouraged to answer as many as questions as they could at each pause, and provide their best guess if they did not know the answer. Approximately ten questions were to be expected during each three-minute pause. After the subjects were familiarized with the MAT system, the NASA-TLX, and the SAGAT procedure, they performed a 5-minute practice session. A longer practice session could be performed upon individual request.

After a subject felt comfortable with the task to be performed, the actual situation awareness experiment began. The subject monitored the green light, the red light, the scales, and tried to keep the amount of fuel in tanks A, B, C, and D at the desired levels. Three to six four-minute freezes randomly occurred during the simulation. The subject spent the first minute of the freeze rating their perceived workload using the NASA-TLX rating scale. The three remaining minutes were used to complete SAGAT queries. The subject answered as many questions as possible, guessing if necessary. After four minutes, the simulation resumed. The participant then took a short break before performing the battery of psychological tests.

The second testing session began with a five-minute review and practice. After a warm up period, the participant performed another trial of the MAT battery task, followed by a short break, and then the remaining psychological tests were conducted.

Operator performances were collected as reaction times to the lights and scales, proportions of light and scale misses, and the average deviations of fuel levels from the target levels. Operator situation awareness was measured by the SAGAT method.

Results

A multiple regression mixed model was created for each of the response variables, using the SAS statistical program. Treatment factors were included as predictor variables, with subjects as random block effects.

Effects of task factors on task performance

Table 1 shows the effects of automation, pace and task duration on operators' performance, as measured by 10 different response variables. Automation showed a positive effect on monitoring task performance. Increased automation resulted in shorter reaction times to the green light and the scales, and lower proportions of green light and scales misses. However, increased automation had a negative effect on the performance of the resource management task, as the operator became less effective in keeping the amount of fuel in tanks B, C, and D at the desired levels. Overall, automation improved performance of the monitoring task, yet decreased performance of the resource management task.

System pace yielded mixed effects on monitoring task performance. At faster paces, the operator had shorter reaction time times to the green light, the red light, and the scales, yet higher proportions of green light, red light, and scales misses. Faster pace yielded a negative effect on the resource management task. The operator had larger deviations of the fuel levels in three out of four tanks when the system was at a faster pace.

Table 1. The significant effects of task factors on task performance (p < .05)

Task Factors	**Task Performance**									
	Reaction times			**Misses**			**Deviations**			
	S	R	G	MS	MR	MG	D_A	D_B	D_C	D_D
Automation	50.27 * <.0001**		6.08 * 0.0153**	41.93 * <.0001**		5.12 * 0.0260**		4.43 * 0.0375**	4.91 * 0.0287**	14.68 * 0.0002**
Pace	157.52 * <.0001**	14.90 * 0.0002**	76.66* <.0001**	110.41 * <.0001**	10.31* 0.0017**	32.60 * <.0001**	30.12 * <.0001**	67.30 * <.0001**	41.94 * <.0001**	23.75 * <.0001**
Duration	6.88 * 0.0102**	11.52 * 0.0010**								6.20 * 0.0144**
Automation x Pace	15.01 * 0.0002**								5.23 * 0.0236**	
Automation x Duration									9.81 * 0.0021**	22.48 * <.0001**
Pace x Duration									19.29 * <.0001**	18.44 * <.0001**
Automation x Pace x Duration										

** F value*

*** Pr > F*

S = Scales, R = Red light, G = Green light, MS = Scale misses, MR = Red light misses, MG = Green light misses, and D_I *= The deviation of fuel in tank I from its target level*

Table 2. The significant effects of task factors on situation awareness ($p < .05$)

Task Factors	**Situation Awareness**						
	Lights and Scales			**Pumps and Fuel Levels**			
	L1	L2	S	P_fail	P_sup	Level	P_plans
Automation	6.24 * 0.0140**			14.33 * 0.0003**	8.02 * 0.0056**	80.67 * <.0001**	8.19 * 0.0050**
Pace		9.96 * 0.0021**			13.17 * 0.0005**	9.64 * 0.0024**	
Duration							
Automation x Pace					7.99 * 0.0054**	4.93 * 0.0279**	
Automation x Duration							
Pace x Duration							
Automation x Pace x Duration							

* *F value*

** *Pr > F*

(L1) the current status of the lights (on/off), (L2) if the lights were currently in the desired conditions, (S) if the scales were currently in the desired conditions, (Level) if the amounts in tanks were in the desired conditions, (P_fail) the current status of the pumps (i.e. malfunctioning, operating, and not operating), (P_sup) the current task of the pumps (i.e. supplying which tanks), and (P_plans) future plans for the pumps (which pumps to turn on)

Longer task duration caused a poorer monitoring task performance, such as longer reaction times to the red light and the scales. However, longer task duration had positive effects on the resource management task performance. The operator could maintain the level of fuel in tank D closer to the target level with the extended task time.

The interactions between automation and task pace, between automation and task duration, and between pace and duration were found to significantly affect task performance (i.e. reaction time to the scales, and deviation of the fuel levels in tanks C and D).

Effects of system factors on situation awareness

Table 2 shows the effects of automation, pace and task duration on operators' situation awareness, measured with seven different response variables (SAGAT scores). Level of automation yielded negative effects on situation awareness. The operator had less knowledge regarding the light status, the pumps, fuel levels, and future pump plan while the system was on auto mode. Faster pace resulted in better knowledge regarding the pumps and the fuel levels, but not the lights. Task duration had no effect on situation awareness. The interaction between automation and task pace was found to significantly affect situation awareness regarding the current task of the pump (P_sup) and the levels of fuel in tanks (Level).

Discussion

Automation yielded better monitoring task performance, yet poorer resource management task performance. Operators may have over trusted the automated resource system, and solely focused their attention on the monitoring task (Endsely, 2000). Furthermore, the automated system in this study allowed the operator to remove him/herself completely out of the resource management sub-system. The operator therefore may not have realized the current status of the resource management sub-system, resulting in out of the loop performance decrements (Endsley, 2000; Endsley and Kaber, 1999).

A similar result was obtained for operator situation awareness regarding the resource management sub-system. The operator had lower situation awareness regarding the resource management sub-system when a higher level of automation was applied. The previous hypotheses to explain operator performance are applicable to operator situation awareness as well.

As automation is positively correlated with monitoring task performance, it is reasonable to expect that automation would boost situation awareness regarding the monitoring sub-system. Surprisingly, the situation awareness regarding the status of the light was lower when the automation level was at a higher level. This unexpected result may be affected by the situation awareness measurement method used. YES/NO type questions were administered to measure the situation awareness on the status of the light. Using this type of question, the possibility that the operator would accidentally select the wrong answer could be as high as 50%.

Increasing the dynamics of the system may only deteriorate operator performance in some aspects. As the pace of the monitoring task increased, the operator had a greater numbers of misses, yet responded to the system faults faster. Higher frequency of the faults occurred during the faster pace condition, making the operator more alert to the changes, and so could more quickly react to such changes. However, shorter timeout periods during the faster pace could have prohibited the operator from detecting the faults, resulting in the greater proportion of misses.

The resource management task, which is combined monitoring skills and strategic planing skills, deteriorated as the task pace increased. A limited amount of time prevented the operator from gathering information, planing to achieve the goal, and/or performing the necessary actions. Therefore, a larger deviation of fuel in the tanks from the desired levels resulted.

In addition to the limited amount of time available, an alternate explanation regarding attention allocation may be forwarded. Several factors, such as movement characteristics, object salience, shape, or color could affect operator attention allocation (Endsley, 1988). Due to the highly dynamic and simplistic monitoring sub-task, the operator may have focused attention and effort more intently on the monitoring task rather than the resource management task. As a result, poor resource management task performance was obtained.

Unlike the resource management task performance, increasing the task pace may actually improve operator situation awareness. The operator had greater situation awareness regarding pump supplies and levels of the fuels in the faster pace environment. As stated earlier, the operator may dedicate him/herself to responding to the faults from the monitoring sub-system. Therefore, the operator may have no time to develop strategies, and/or perform the actions to accomplish the resource management goal even though high situation awareness of this sub-system was achieved. This scenario agrees with Endsley (1996, in Thomas and Charlton, 1996, Chapter 9) in that high situation awareness may not always yield superior performance. Factors such as the ability to develop the strategic planning must be achieved to attain good performance.

The monitoring task performance degraded after performing the task for a lengthy amount of time. It could be summarized that fatigue and boredom were the causes leading to poorer performance. However, fatigue and boredom from long task duration did not have any effect on resource management task performance or on situation awareness regarding any system elements. The operator may have gained some experience and, therefore, could develop strategies for resource management. Experiencing with SAGAT freezes may also gave the operator clues of what to expect (Sarter and Woods, 1995), so could answer the questions regardless of fatigue or boredom that may have occurred.

In summary, task factors had significant effects on performance and situation awareness of novice operators of a simulated complex system. Automation level and pace of work showed a stronger affect on task performance than task duration did. Automation was the strongest factor affecting situation awareness, while pace yielded lesser effects. No effect of task duration on situation awareness was found.

Future investigation is required to confirm the current findings, as well as to provide opportunity for further discovery. Possible studies to be conducted include incorporating different forms of task settings such as automation failure warning light, level of severity if failing to meet the goals, or compensation levels into the simulated task used in this study; incorporating different interface designs; conducting a long term study; or having operators work in teams.

References

Comstock, J.R., Jr., and Arnegard, R.J. (1992). *The Multi-variable Attribute Task battery for Human Operation workload and Strategic Behavior Research.*

Endsley, M.R. (1988). Design and Evaluation for Situation Awareness Enhancement, *Proceedings of the Human Factors Society 32th Annual Meeting,* 1, 97-101.

Endsley, M.R. (1995). Toward the Theory of Situation Awareness in Dynamic Systems, *Human Factors, 37(1),* 32-64.

Endsley, M.R., (2000). Situation Awareness, Automation & Free Flight, *http://vega3.uneec.eurocontrol.fr/endsley.htm*, Date accessed: 11/29/2000.

Endsley, M.R., and Kaber, D. B. (1999). Level of Automation Effects on Performance, Situation awareness and Workload in a Dynamic Control Task. *Ergonimics*, **42**, 462-492.

Hancock, P.A., and Caird, J.K. (1993). Experimental Evaluation of a Model of Mental Workload. *Human Factors*, 35(**3**), 413-429.

Hanky, M.J., Dingus, T.A. (1990). A Validation of SWAT as a Measure of Workload Induced by Changes in Operator Capacity. In *Proceedings of the Human Factors Society 34th Annual Meeting,* (pp. 112-115). Santa Monica, CA: Human Factors and Ergonomics Society.

O'Brien, T.G., and Charlton, S.G. (1996). *Handbook of Human Factors Testing and Evaluation*, Lawrence Erlbaum Associates. New Jersey.

Raby, M., and Wickens, C.D. (1994). Strategic Workload Management and Decision Biases in Aviation. *The International Journal of Aviation Psychology*, 4(*3*), 211-240.

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.

Wickens C.D., and Hollands, J.G. (2000). *Engineering Psychology and human Performance* 3rd Ed, Prentice-Hall Inc. New Jersey.

Chapter 5. Task Performance and Situation Awareness: The Role of Attention Limitations

A paper to be submitted to *Ergonomics*

Jaruwan Klamklay and Patrick E Patterson

Abstract

This study examined the effects of human attention limitations on situation awareness and task performance of operators performing a simulated complex task. A simplified pilot task required operators to monitor system status and manage system fuel resources. Eighty participants performed the simulated task under different task paces, duration, and automation levels. Speed and accuracy were used to measure task performance and the Situation Awareness Global Assessment Technique (SAGAT) was used to quantify situation awareness. The subjects also performed five psychological tests to measure five attention limitations. Associations between task performance and each of five attention limitations, and between situation awareness each of five attention limitations, were identified using Pearson's correlation coefficients. The results showed that operators with high scores on four of the five attention tests had high task performance, and high scores on three of the five tests were associated with high situation awareness. The regression analysis suggested that the scores from the Visual search and the Simon effect tests could predict task performance and situation awareness of operators of a simulated complex system.

Introduction

Situation awareness (SA) is the knowledge of the current state of a system, the understanding of what is going on in the system at a certain instant of time, and what will happen in the near future. Endsley (2000) defined situation awareness as simply "knowing what is going on around you".

Situation awareness is needed for operators to perform tasks effectively (Endsley 2000). To reach a goal, the operator must begin by forming situation awareness, make a decision, and then performing necessary actions. These three steps occur continuously and are highly interrelated, therefore the quality of each step depends on the previous one (Endsley,1988; Endsley, 1995).

The importance of situation awareness becomes greater as the complexity and dynamic nature of a system increases. Due to the interconnection and interdependence among system components, a single component error or failure could result in numerous side effects. Additionally, an individual's current situation awareness affects the way the new information is perceived and interpreted (Sandom, 1999 & Endsley, 1988). As a result, incomplete or inaccurate current situation awareness at an instant in time will lead to poorer situation awareness at later time.

Operators of a complex system must continuously keep track of what is transpiring, anticipating potential problems, and preparing to solve any problems encountered. When a problem does occur, operators need to quickly make decisions and perform corrective actions to bring the system back to normal conditions. Incomplete or inaccurate prior SA will require more time to revise current situation awareness, to make a decision, and perform necessary actions. The worst case scenario occurs when operators cannot successfully perform all corrective actions within given time limitations, which can lead to disastrous consequences.

Attention resources are critical elements in the information processing process. It is the first step that selectively focuses on some stimuli and while ignoring others (Coglab, 2000a) to gather the information needed to form SA. While numerous system parameters are competing for the operators' attention, a person's ability to accurately perceive multiple items simultaneously is limited. Operators, therefore, overcome the limitation by sampling to attend to only some of the information considered to be important in attaining the goal. The higher the degree of contribution to goal success in the operator's opinion, the more attention is allocated (Fracker,1989). For example, fighter pilots directed more attention to the targets considered more critical to mission success (Endsley and Smith, 1996); air traffic controllers paid more attention to aircraft rather than other less important information (Endsley and Rogers,1998); and drivers paid more attention to cars closer to them rather than ones further away (Gugerty, 1998).

However, overcoming attention limitations by sampling to attend to selected information leads to SA error. Jones and Endsley (1996) found that about 35% of SA error occurred because the operators failed to attend to the necessary information, which was present. Schutte and Trujillo (1996) found that the pilots who paid the most attention to the flight rather than the system, failed to properly detect and manage a fuel leak situation, placing their airplanes in critical conditions.

Endsley and Bolstad (1994) investigated whether individual differences such as spatial ability, attention, perception, memory, and cognitive functions, could predict a pilot's situation awareness. Twenty-five former military fighter pilots in groups of six (some subjects participated on more than one team) performed air-to-air engagement in a real time simulator. A subject's situation awareness was measured using Situation Awareness Global Assessment Technique (SAGAT). Nineteen military attribute tests used in military selection batteries were then administered to the subjects to measure situation awareness correlates, by determining Pearson correlation between SA and the 31 variables. The reaction times from the attention sharing tests moderately correlated with pilot SA (rs = -.138, and -.25). However, tracking performance from the attention sharing test was highly correlated with SA (r = .717).

Carretta, Perry,& Ree (1996) conducted an experiment using 171 USAF pilots who performed an air-to-air mission exercise whose SA was collected by using a supervisor and peer (pilots rate other pilots) rating methods. The participant's differences were measured for cognitive ability, psychomotor ability, and personality. When holding flying experience constant, they found that general cognitive ability based on working memory, spatial reasoning, and divided attention were significant predictors of SA (r = .70).

O'Hare (1997) administered the WOMBAT Situation Awareness and Stress Tolerance Test and the Walter Reed Performance Assessment Battery to a group of 24 males having differing ages, occupations, and computer related experiences. WOMBAT is a simulation of a complex-multiple task requiring a high level of SA. The Walter Reed Performance Assessment Battery includes tests for cognitive abilities- Pattern Recognition, Digit Recall, Six-letter search, and Manikin. The six-letter search measured operator selective attention. No correlation was found between the six-letter search performance and the WOMBAT performance.

While several studies confirmed the important role of attention in task performance and situation awareness, correlations between attention tests and task performance or situation awareness were confounded. It could be hypothesized that attention measurements caused such inconsistency. The correlations between attention and task performance and attention and situation awareness could be more distinct if basic attention limitation measurements were used. Below are details regarding natural limitations of attention found in previous psychological studies, and the tests developed to demonstrate such limitations.

First, knowledge of where targets will occur could reduce response time. Posner et al (1978) found that a spatial cue of a target would help individuals by directing their attention to the target so they could respond to the target faster. One test used to demonstrate this phenomena is called Spatial Cuing.

Second, an individual's attention resource is limited and it may process a limited amount of information at any given time (Raymond, Shapiro, & Arnell, 1992). A new item will be processed only if the pervious information has finished being processed. Therefore, when items are present in series, the next item will be processed only if the previous item has been processed. A test used to demonstrate this phenomenon is called Attentional Blink.

Third, visual search time depends on the relationship between the targets and their background. Visual search time is generally shorter when targets are more distinguishable. A test used to demonstrate this phenomenon is called Visual Search.

Fourth, automatized responses that conflict with the desired behavior will slow down information processing. Therefore, humans will be slow in responding to a target that is contrary to automatic behavior, although it might require a simple response. One test used to demonstrate this phenomenon is called Stroop Effect.

Lastly, location irrelevancy between the target and the response key may slow down the response identification process (Simon 1969). A test used to demonstrate this phenomena is called Simon Effect.

These five attention limitations may directly affect operator performance and situation awareness. To date, no studies have examined their relationships with operator performance or situation awareness.

The first objective of this study was to determine the relationships of the five attention limitations with operator performance and situation awareness. The following hypotheses were investigated: (1) higher scores on the Spatial Cuing test will be positively correlated with better task performance and higher situation awareness, (2) higher scores on the Attentional Blink test will be positively correlated with better task performance and higher situation awareness, (3) higher scores on the Visual Search test will be positively correlated with better task performance and higher situation awareness, (4) higher scores on the Stroop Effect test will be positively correlated with better task performance and higher situation awareness, and (5) higher scores on the Simon Effect test will be positively correlated with better task performance and higher situation awareness.

The second objective was to determine the effects of attention limitations on task performance and situation awareness of operators performing the MAT task under different conditions of automation level, pace, and work duration.

Methods

Participants

Eighty undergraduate engineering students (25 female and 55 males) ages between 19-25 years from various disciplines were enrolled on a voluntary basis. The study was evaluated and approved by the Iowa State University Human Subject Committee.

Tasks

The operator was required to perform two types of tasks; the Multi-variable Attribute Task battery for Human Operation workload and Strategic Behavior Research (MAT) simulation task, and five psychological tests.

The MAT simulation task

This study used the Multi-variable Attribute Task battery for Human Operation Workload and Strategic Behavior Research (Comstock and Arnegard, 1992) to simulate a simple complex system. Each operator of the system was required to perform two sub-tasks. The first sub-task was system monitoring, in which the operator monitored two lights and four vertical scales for signs of system abnormalities. Under normal conditions, the left light would always be green. However, if a system abnormality occurred, the green light would go out. The operator responded to the disappearance of the green light by pressing the assigned key as soon as possible. The right light was normally off; however, if a system abnormality occurred a red light would turn on in that position. Similarly, once the operator noted the red light was on, he/she pressed the assigned key as soon as possible. The operator would see feedback as the green light would immediately turn back on, and the red light would immediately turn off.

In addition to the lights, the operator also monitored the vertical scales below the lights. Each scale had a yellow pointer that could fluctuate one unit below to one unit above the centerline. If a system fault occurred, the corresponding scale would shift out of the normal range. Regardless of the abnormal shifting direction, the operator could correct the fault by pressing the assigned key. The feedback to a correct response was given by the presence of a yellow bar at the bottom of the scale that indicated an out of range condition, and a return to the center of that scale pointer. If the operator did not notice the fault events (i.e. the lights, the scales), the situation would return to normal condition after a selected time-out period.

The second sub-task was resource management. This sub-system contained six fuel tanks connected by pipelines. The main tanks were tanks A and B, each having 4000-unit capacity. The supply tanks C and D contained a maximum of 2000 units each, and two additional unnamed supply tanks had unlimited capacities. The system used the fuel from tanks A and B, so fuel levels in these two tanks continuously decreased while the system was operating. The operator's task was to maintain the fuel in tanks A and B at their specific levels, as well as trying to keep tanks C and D full. Both numbers underneath tanks A, B, C, and D, and the green levels in the tanks, represented the current amount of fuel in the tanks. To meet the goal, the operator used the eight pumps provided to transfer fuel between tanks. The pumps transferred fuel from one tank to others in the directions indicated by arrows on the display. Keys numbered 1 to 8 were used as toggle switches to turn the corresponding pumps On/Off. No interactions from the operator were required if resource management was in its automatic mode.

The experimenter could set several system parameters including the disappearance of the green light, the appearance of the red light, the out-of-range fluctuation of the scales, time-out periods, resource management automatic mode, pump failures, service time to the failed pumps, pump flow rates, and the fuel consumption rates of tanks A and tank B.

Task performance was measured from the average reaction times to detect and correct system malfunctions (i.e. the scales, the green light, and the red light), proportions of misses (i.e. number of uncorrected malfunctions/ total number of malfunctions), and the average deviation of fuel levels of tanks A, B, C, and D from their target levels.

Each operator's SAGAT score was calculated by comparing the operator's answers to what actually happened. Each query was graded independently. Seven groups of questions presented to the subjects were: (1) the current status of the lights (on/off), (2) if the lights were currently in the desired conditions, (3) if the scales were currently in the desired conditions, (4) if the fuel amounts in the tanks were in the desired ranges, (5) the current status of the pumps (i.e. malfunctioning, operating, and not operating), (6) the current tasks of the pumps (i.e. supplying which tanks), and (7) future plans for the pumps (which pumps to turn on).

Questions were either Yes/No questions, or multiple answer questions. For Yes/No questions, if the question was answered correctly the score would be one, otherwise zero. For the questions that contain more than one possible answer, the score was determined by the number of correct answers divided by the summation of the number of correct and incorrect answers. Possible answers that were not chosen were considered incorrect answers. The scores for this type of question ranged between zero and one.

Psychological tests

Five classic psychological tests were administered to the participant as to measure attention limitations.

Attentional Blink. Each subject watched several series of ten letters, in which the newest letter overwrote the previous letter. Each series contained the first target letter, and/or the second target letter. The separation between the presence of the first target and the second target may vary between series. Each subject reported the presence of both target letters. The average percentage of time a subject could correctly identify the second targets was recorded, with greater values indicating better use of attention resources.

Spatial Cuing. Each subject was required to respond to the presence of targets as quickly as possible under three conditions: (1) when a cue correctly identified a location of a target, (2) when a cue incorrectly identified a location of a target, and (3) when there was no cue provided. The average reaction times under these three conditions were recorded with a faster reaction time indicating better use of attention resources.

Visual Search. Each subject performed two types of visual search tasks. The first type was a conjunctive search, in which the targets shared at least one characteristic (i.e. color, or shape) with the background. The second type was a feature search, in which the targets were completely different form the background and, therefore, would "pop-out" of the display. For the conjunctive search, each subject determined whether a green circle was presented among green squares and blue circles. For the feature search, each subject determined whether a green circle was presented among blue squares. The non-target items, such as green squares and blue circles, were distracters. Average reaction times were recorded with faster visual search times indicating better use of attention resources.

Stroop Effect. Each subject was required to identify the font colors of the words "BLUE", "RED", and "GREEN". The words could be written in the font colors blue, red, or green. The average reaction times to identify the font color when the font color and the word name were the same, and when they were different, were recorded with a faster reaction time indicating better use of attention resources.

Simon Effect. Each subject was required to respond to the presence of the target as quickly possible under four different conditions: (1) when the target was presented on the right, and the response key was located on the right, (2) when the target was presented on the left, and the response key was located on the left, (3) when the target was presented on the right, and the response key was located on the left, and (4) when the target was presented on the left, and the response key was located on the right. Average reaction times and average proportion of errors from conditions 1 to 4 were recorded with faster reaction times, and fewer errors indicating better use of attention resources.

Experimental design

The MAT task contained 2 levels of pace, 2 levels of task duration, and 2 levels of automation, creating 8 possible treatments. Each subject was assigned to perform two different simulation treatments, using balanced incomplete block design. To prevent the participant from guessing the nature of the simulation, each of the two treatments was randomly chosen to perform on different days with at least three days between treatments.

Five attention tests were administered to each of the subjects in random order. The dependent variables were the operators' MAT task performance and MAT situation awareness. The predictor variables were pace, automation, task duration, and the scores from the five attention tests.

Procedure

The first session began with a 15-minute introduction and practice, allowing the participants to become familiar with the MAT program, the NASA-TLX workload rating scale, and the SAGAT measurement. MAT system components and their fundamental characteristics were explained. The subjects were informed in advance to expect system failure events such as the disappearance of the green light, the appearance of the red light, the deviations of the scales, pump failures, and automation mode failures. Each subject practiced responding to the event faults and controlling the pumps via the computer keyboard. Next, the NASA-TLX self-rating procedure was introduced including the meaning of each sub-scale. Each subject learned how to use the NASA-TLX scale, followed by a practice session. The SAGAT procedure was then introduced to the subjects. The subjects were informed to expect freezes and questions during the experimental session. The subjects were encouraged to answer as many as questions as they could at each pause, and provide their best guess if they did not know the answer. Approximately ten questions were to be expected during each three-minute pause. After the subjects were familiarized with the MAT system, the NASA-TLX, and the SAGAT procedure, they performed a 5-minute practice session. A longer practice session could be performed upon individual request.

After a subject felt comfortable with the task to be performed, the actual situation awareness experiment began. The subject monitored the green light, the red light, the scales, and tried to keep the amount of fuel in tanks A, B, C, and D at the desired levels. Three to six four-minute freezes randomly occurred during the simulation. The subject spent the first minute of the freeze rating their perceived workload using the NASA-TLX rating scale. The three remaining minutes were used to complete SAGAT queries. The subject answered as many questions as possible, guessing if necessary. After four minutes, the simulation resumed. The participant then took a short break before performing the battery of psychological tests.

The second testing session began with a five-minute review and practice. After a warm up period, the participant performed another trial of the MAT battery task, followed by a short break, and then the remaining psychological tests were conducted.

Operator performances were collected as reaction times to the lights and scales, proportions of light and scale misses, and the average deviations of fuel levels from the target levels. Operator situation awareness was measured by the SAGAT method.

Results

The JMP statistical software (JMP 4.0.2) was used to determine the relationships between the factors of interest, as Pearson correlation coefficients. Subject random block effects may affect the Pearson correlation coefficients so only the data from one treatment was randomly chosen to be included in the data analysis.

Relationships between attention limitations and task performance

The results (Table 1) show that higher scores on the Spatial Cuing test were positively associated with better task performance. The operator with faster reorienting times tended to have fewer scale misses, fewer red light misses, and less fuel deviations in tanks A and B.

Attentional Blink test scores yielded a positive correlation with the fuel deviation in tank A. Operators with faster information processing rates tended to have larger fuel deviation.

Visual conjunctive search times were negatively associated with proportion of green light misses. Operators who were quicker in detecting conjunctive targets tended to have more green light misses.

Visual feature search times were positively associated with red light misses, scales misses, and fuel deviations. Operators who were quicker in detecting the feature target tended to have fewer red light misses, fewer scales misses, and less fuel deviations in tanks A, B, and C.

Higher Stroop Effect test scores were positively correlated with better task performance. Operator who could quickly identify the required response when facing the situation that counter the automatic behavior tended to be quicker in responding to the red light, and better in maintaining the fuel in tanks B and C closer to the desired levels.

Higher Simon Effect test scores were positively correlated with better task performance. The operator who could quickly respond to the target regardless whether the location of the stimuli matched the location of the response tended to have shorter reaction time to the red light, fewer scales misses, fewer green light misses, and less fuel deviations in tanks A, B, and D.

Relationships between attention limitations on situation awareness

The results (Table 2) show that higher scores on the Spatial Cuing test were positively correlated with higher situation awareness. Operators with faster attention reorientation tended to have more accurate knowledge regarding the current status of the lights.

Higher scores on the Attentional Blink test were negatively correlated with higher situation awareness. Operators with faster information processing rates tended to have less accurate knowledge whether the lights were in the desired conditions.

Higher scores on the Visual Search test (feature search) were positively correlated with higher situation awareness. The operators who were quicker in identifying feature targets tended to have more accurate situation awareness regarding the currents status of the pumps.

Higher scores on the Simon Effect test were positively correlated with higher situation awareness. Operators with faster and more accurate responses, when the location of the stimuli didn't match with the location of the response, tended to have more accurate knowledge regarding the current status of the pumps.

Effects of attention limitations on task performance and situation awareness

A multiple regression mixed model was created for each of the response variables, using the SAS statistical program. Measurements from each of the attention tests and treatment factors were included as predictor variables, with subjects as random block effects. All 160 data points were included in the analysis. Table 3 shows a matrix of correlation coefficients between each pair of the attention test scores.

Effects of attention limitations on task performance

The results (Table 4) show that performance of each of the five attention tests significantly affected MAT task performance.

Effects of attention limitations on situation awareness

The results (Table 5) show that performance of each of the Visual search and the Simon effect tests significantly affected MAT situation awareness.

Table 1. The significant correlations between attention test scores and task performance ($p < .05$)

Test Scores		Task Performance									
		Reaction times			Misses			Deviations			
		S	R	G	MS	MR	MG	D_A	D_B	D_C	D_D
Spatial Cuing	SCV							.2839	.3360		
	SCIV				.3003	.5033		.4095	.3101		
Attentional Blink	AB0							.2243			
	AB2							.2688			
	AB4							.2590			
Visual Search	VSCA4						-.2446				
	VSCA64						-.2626				
	VSFA16							.3135	.2367	.2457	
	VSFA64							.3663			
	VSFP16				.2641	.3305			.2417		
Stroop Effect	Ssame								.3196	.2450	
	Sdiffer		.2906								
Simon Effect	SimCL		.3154								
	SimCR						.2200				
	SimICL				.2471		.2461		.2527		
	SimICR								.2221		
	SimICLe										.2752
	SimICRe							.3822	.3445		

S = Scales, R = Red light, G = Green light, MS = Scale misses, MR = Red light misses,
MG = Green light misses, and D_I = The deviation of fuel in tank I from its target level
SCV = Spatial Cuing when the cue is valid, SCIV = Spatial Cuing when the cue is invalid,
ABi = Attentional Blink when the separation between target is 100i milliseconds
VSCAi / VSCPi = Visual Search Conjunctive Absent / Present with i distracters
VSFAi / VSFPi = Visual Search Feature Absent / Present with i distracters
Ssame / Sdiffer = Stroop Effect when font color and font name are the same / different
SimCR / SimCL = Simon Effect when the stimuli is on the right / left and the response key is on the right / left
SimICR / SimICL = Simon Effect when the stimuli is on the right / left and the response key is on the left / right

Table 2. The significant correlations between attention test scores and situation awareness (p < .05)

Test Scores		Situation Awareness						
		Lights and Scales			Pumps and Fuel Levels			
		L1	L2	S	P_fail	P_sup	Level	P_plans
Spatial Cuing	SCV	-.2968						
Attentional Blink	ABS2		-.2986					
Visual Search	VSFA16				-.2859			
Stroop Effect	Ssame		-.2483	-.3472				
Simon Effect	SimICRe				-.2455			

(L1) the current status of the lights (on/off), (L2) if the lights were currently in the desired conditions, (S) if the scales were currently in the desired conditions, (Level) if the amounts in tanks were in the desired conditions, (P_fail) the current status of the pumps (i.e. malfunctioning, operating, and not operating), (P_sup) the current task of the pumps (i.e. supplying which tanks), and (P_plans) future plans for the pumps (which pumps to turn on)

SCV = Spatial Cuing when the cue is valid, SCIV = Spatial Cuing when the cue is invalid,
ABi = Attentional Blink when the separation between target is 100i milliseconds
VSCAi / VSCPi = Visual Search Conjunctive Absent / Present with i distracters
VSFAi / VSFPi = Visual Search Feature Absent / Present with i distracters
Ssame / Sdiffer = Stroop Effect when font color and font name are the same / different
SimCR / SimCL = Simon Effect when the stimuli is on the right / left and the response key is on the right / left
SimICR / SimICL = Simon Effect when the stimuli is on the right / left and the response key is on the left / right

Table 3. A matrix of correlation coefficients between each pair of the attention test scores

	SCN	SCV	SCIV	AB0	AB2	AB4	AB6	AB8	VSCA4	VSCA16	VSCA64	VSCP4	VSCP16	VSCP64	VSFA4
SCN	1	0.073	0.054	0.099	-0.031	-0.277	0.0295	0.053	0.0529	0.0176	-0.1628	0.0411	0.0498	0.0263	0.177
SCV	0.073	1	0.028	-0.029	-0.043	0.196	0.2964	0.339	0.1772	0.206	0.144	0.1635	0.0531	0.1971	0.1406
SCIV	0.054	0.028	1	0.052	-0.069	0.03	0.0624	0.083	-0.0105	-0.0813	-0.0765	-0.018	-0.0455	-0.0506	0.2719
AB0	0.099	-0.029	0.052	1	0.26	0.153	0.2311	0.098	-0.1073	-0.0914	-0.2021	0.0106	-0.0344	-0.0691	0.0115
AB2	-0.031	-0.043	-0.07	0.26	1	0.259	0.2849	0.294	0.0173	0.0067	0.0585	-0.11	0.2138	0.0166	-0.0679
AB4	-0.277	0.196	0.03	0.153	0.259	1	0.6597	0.558	0.0006	-0.003	0.1121	-0.056	-0.0811	-0.0638	0.0445
AB6	0.03	0.296	0.062	0.231	0.285	0.66	1	0.762	0.034	0.1221	0.1082	-0.109	0.0346	0.09	0.0953
AB8	0.053	0.339	0.083	0.098	0.294	0.558	0.7615	1	0.1656	0.0946	0.1477	-0.008	0.0458	0.1403	0.0765
VSCA4	0.053	0.177	-0.01	-0.107	0.017	6E-04	0.034	0.166	1	0.4952	0.5531	0.3089	0.4074	0.5297	0.3327
VSCA16	0.018	0.206	-0.08	-0.091	0.007	-0.003	0.1221	0.095	0.4952	1	0.6494	0.1535	0.3175	0.5333	0.4314
VSCA64	-0.163	0.144	-0.08	-0.202	0.059	0.112	0.1082	0.148	0.5531	0.6494	1	0.1519	0.3473	0.6756	0.3794
VSCP4	0.041	0.164	-0.02	0.011	-0.11	-0.056	-0.109	-0.008	0.3089	0.1535	0.1519	1	0.0929	0.0826	0.2463
VSCP16	0.05	0.053	-0.05	-0.034	0.214	-0.081	0.0346	0.046	0.4074	0.3175	0.3473	0.0929	1	0.4186	0.2582
VSCP64	0.026	0.197	-0.05	-0.069	0.017	-0.064	0.09	0.14	0.5297	0.5333	0.6756	0.0826	0.4186	1	0.2133
VSFA4	0.177	0.141	0.272	0.012	-0.068	0.045	0.0953	0.077	0.3327	0.4314	0.3794	0.2463	0.2582	0.2133	1
VSFA16	0.169	0.073	0.027	-0.038	-0.066	-0.08	0.0009	0.119	0.2275	0.1052	0.2822	0.3	0.1767	0.1605	0.2664
VSFA64	0.13	0.086	-0.09	-0.096	-0.198	-0.103	-0.206	-0.166	0.3332	0.2861	0.362	0.337	0.2187	0.2999	0.4982
VSFP4	0.216	0.12	-0.1	-0.197	-0.165	-0.359	-0.306	-0.135	0.2643	0.1885	0.1728	0.2823	0.2495	0.2195	0.3395
VSFP16	0.161	0.012	0.734	0.068	-0.12	-0.041	0.0537	0.119	0.1248	-0.014	-0.008	0.0918	0.0892	0.0337	0.3907
VSFP64	0.105	-0.002	-0.05	-0.057	-0.009	-0.132	-0.214	-0.083	0.362	0.1805	0.1914	0.3204	0.3346	0.2052	0.2826
Ssame	0.149	0.1	-0.01	0.207	-0.102	-0.099	-0.104	-0.171	0.24	0.1926	0.1543	0.1837	0.0915	0.1835	0.1713
Sdiffer	0.191	0.289	0.017	8E-04	-0.063	-0.123	-0.117	-0.069	0.2949	0.3037	0.2504	0.2473	0.1658	0.2843	0.3303
SimCL	0.136	0.23	-0.02	-0.051	-0.087	-0.1	-0.06	-0.091	0.3108	0.2274	0.1528	0.3518	0.3227	0.2213	0.2848
SimCR	0.122	0.386	-0.05	-0.071	-0.083	-0.084	-0.03	-0.045	0.3166	0.3231	0.2269	0.27	0.3116	0.2513	0.3788
SimICL	0.17	0.175	-0.05	0.137	-0.056	-0.099	-0.101	-0.185	0.1991	0.0957	0.0019	0.2339	0.158	0.0794	0.2565
SimICR	0.148	0.261	-0.09	0.065	-0.036	-0.073	-0.072	-0.031	0.3848	0.2945	0.1016	0.3348	0.3188	0.2061	0.314
SimCLe	-0.087	-0.151	0.16	-0.029	0.17	0.108	0.0723	0.148	-0.1824	-0.0102	0.0734	-0.147	0.0012	-0.1525	0.0704
SimCRe	-0.219	-0.139	-0.1	0.05	0.072	0.048	0.0589	0.049	-0.1468	0.0317	-0.0244	-0.214	-0.1328	-0.0915	-0.1722
SimICLe	-0.127	-0.116	0.067	0.29	0.165	0.218	0.1523	0.076	-0.3511	-0.2221	-0.3005	-0.144	-0.032	-0.2307	-0.0436
SimICRe	-0.011	-0.109	0.268	0.176	0.015	0.08	0.1289	0.172	-0.3057	-0.2302	-0.3339	-0.003	-0.1384	-0.3441	-0.1567

Table 3 (continued). A matrix of correlation coefficients between each pair of the attention test scores

	VSFA16	VSFA64	VSFP4	VSFP16	VSFP64	Ssame	Sdiffer	SimCL	SimCR	SimICL	SimICR	SimCLe	SimCRe	SimICLe	SimICRe
SCN	0.1689	0.1298	0.2163	0.1611	0.105	0.149	0.1907	0.1363	0.1215	0.1695	0.1483	-0.0871	-0.2189	-0.127	-0.0111
SCV	0.0734	0.0856	0.1198	0.0124	-0.0016	0.1	0.2893	0.2296	0.3859	0.1746	0.2606	-0.1506	-0.1394	-0.1157	-0.109
SCIV	0.0268	-0.0905	-0.102	0.7337	-0.053	-0.012	0.0172	-0.022	-0.048	-0.0484	-0.0885	0.1603	-0.1009	0.0671	0.2678
AB0	-0.0384	-0.0961	-0.1974	0.0675	-0.0574	0.207	0.0008	-0.051	-0.071	0.1369	0.0647	-0.0291	0.0501	0.2898	0.1758
AB2	-0.066	-0.1983	-0.1652	-0.1198	-0.0094	-0.102	-0.063	-0.087	-0.083	-0.0564	-0.036	0.1697	0.072	0.1649	0.0153
AB4	-0.0798	-0.1027	-0.3585	-0.0407	-0.1323	-0.099	-0.123	-0.1	-0.084	-0.0989	-0.0725	0.1075	0.0481	0.218	0.0796
AB6	0.0009	-0.2061	-0.3055	0.0537	-0.2138	-0.104	-0.117	-0.06	-0.03	-0.1006	-0.0718	0.0723	0.0589	0.1523	0.1289
AB8	0.1187	-0.1657	-0.1346	0.119	-0.0832	-0.171	-0.069	-0.091	-0.045	-0.1849	-0.0308	0.1475	0.0488	0.0757	0.1718
VSCA4	0.2275	0.3332	0.2643	0.1248	0.362	0.24	0.2949	0.3108	0.3166	0.1991	0.3848	-0.1824	-0.1468	-0.3511	-0.3057
VSCA16	0.1052	0.2861	0.1885	-0.014	0.1805	0.193	0.3037	0.2274	0.3231	0.0957	0.2945	-0.0102	0.0317	-0.2221	-0.2302
VSCA64	0.2822	0.362	0.1728	-0.008	0.1914	0.154	0.2504	0.1528	0.2269	0.0019	0.1016	0.0734	-0.0244	-0.3005	-0.3339
VSCP4	0.3	0.337	0.2823	0.0918	0.3204	0.184	0.2473	0.3518	0.27	0.2339	0.3348	-0.1468	-0.2137	-0.1441	-0.0034
VSCP16	0.1767	0.2187	0.2495	0.0892	0.3346	0.092	0.1658	0.3227	0.3116	0.158	0.3188	0.0012	-0.1328	-0.032	-0.1384
VSCP64	0.1605	0.2999	0.2195	0.0337	0.2052	0.184	0.2843	0.2213	0.2513	0.0794	0.2061	-0.1525	-0.0915	-0.2307	-0.3441
VSFA4	0.2664	0.4982	0.3395	0.3907	0.2826	0.171	0.3303	0.2848	0.3788	0.2565	0.314	0.0704	-0.1722	-0.0436	-0.1567
VSFA16	1	0.3049	0.3632	0.2217	0.2548	0.196	0.1307	0.0835	0.1286	0.0926	0.1408	-0.0278	-0.1577	-0.0141	0.0838
VSFA64	0.3049	1	0.3501	0.0761	0.4276	0.207	0.2762	0.2552	0.3112	0.2455	0.203	-0.1488	-0.2429	-0.2719	-0.3275
VSFP4	0.3632	0.3501	1	0.2489	0.4896	0.115	0.2451	0.2948	0.4013	0.2829	0.3256	-0.1311	-0.1354	-0.2124	-0.1022
VSFP16	0.2217	0.0761	0.2489	1	0.256	0.012	0.1257	0.1972	0.143	0.0848	0.1603	0.0565	-0.1151	-0.0791	0.1776
VSFP64	0.2548	0.4276	0.4896	0.256	1	0.138	0.1902	0.2493	0.3584	0.2343	0.3666	-0.086	-0.1344	-0.134	-0.126
Ssame	0.1963	0.2067	0.1154	0.0124	0.1384	1	0.4776	0.2951	0.3789	0.5945	0.3803	-0.0186	-0.1069	-0.0669	-0.1989
Sdiffer	0.1307	0.2762	0.2451	0.1257	0.1902	0.478	1	0.5507	0.5883	0.5024	0.4859	-0.086	-0.1015	-0.1689	-0.1806
SimCL	0.0835	0.2552	0.2948	0.1972	0.2493	0.295	0.5507	1	0.8017	0.7655	0.7484	-0.1242	-0.0683	-0.3213	-0.3234
SimCR	0.1286	0.3112	0.4013	0.143	0.3584	0.379	0.5883	0.8017	1	0.7893	0.7762	-0.1419	-0.0976	-0.3335	-0.2855
SimICL	0.0926	0.2455	0.2829	0.0848	0.2343	0.595	0.5024	0.7655	0.7893	1	0.7618	-0.1579	-0.1901	-0.2527	-0.2456
SimICR	0.1408	0.203	0.3256	0.1603	0.3666	0.38	0.4859	0.7484	0.7762	0.7618	1	-0.1989	-0.145	-0.3276	-0.1753
SimCLe	-0.0278	-0.1488	-0.1311	0.0565	-0.086	-0.019	-0.086	-0.124	-0.142	-0.1579	-0.1989	1	0.4668	0.1155	0.3503
SimCRe	-0.1577	-0.2429	-0.1354	-0.1151	-0.1344	-0.107	-0.102	-0.068	-0.098	-0.1901	-0.145	0.4668	1	0.097	0.3261
SimICLe	-0.0141	-0.2719	-0.2124	-0.0791	-0.134	-0.067	-0.169	-0.321	-0.334	-0.2527	-0.3276	0.1155	0.097	1	0.343
SimICRe	0.0838	-0.3275	-0.1022	0.1776	-0.126	-0.199	-0.181	-0.323	-0.286	-0.2456	-0.1753	0.3503	0.3261	0.343	1

Table 4. The significant effects of attention test scores on task performance (p < .05)

Test scores		**Task Performance**									
		Reaction times			**Misses**			**Deviations**			
		S	R	G	MS	MR	MG	D_A	D_B	D_C	D_D
Spatial Cuing	SCIV					7.95* .0068**		4.99* .0299**			
Attentional Blink	AB8				4.67* .0357**						
Visual Search	VSCA64				4.59* .0371**		5.46* .0235**				
	VSFA16						4.33* .0425**				
Stroop Effect	Ssame								4.58* .0343**	9.68* .0030**	
	Sdiffer									6.13* .0168**	
Simon Effect	SimCL		5.50* .0230**						4.81* .0303**	10.82* .0018**	
	SimICR						6.72* .0125**				
	SimCLe									5.05* .0288**	

** F value*

*** Pr > F*

S = Scales, R = Red light, G = Green light, MS = Scale misses, MR = Red light misses, MG = Green light misses, and
D_I = The deviation of fuel in tank I from its target level
SCV = Spatial Cuing when the cue is valid, SCIV = Spatial Cuing when the cue is invalid,
ABi = Attentional Blink when the separation between target is 100i milliseconds
VSCAi / VSCPi = Visual Search Conjunctive Absent / Present with i distracters
VSFAi / VSFPi = Visual Search Feature Absent / Present with i distracters
Ssame / Sdiffer = Stroop Effect when font color and font name are the same / different
SimCR / SimCL = Simon Effect when the stimuli is on the right / left and the response key is on the right / left
SimICR / SimICL = Simon Effect when the stimuli is on the right / left and the response key is on the left / right

Table 5. The significant effects of attention test scores on situation awareness ($p < .05$)

Test scores		**Situation Awareness**						
		Lights and Scales			**Pumps and Fuel Levels**			
		L1	L2	S	P_fail	P_sup	Level	P_plans
Visual Search	VSCP4				9.14* .0040**			
	VSFA4					4.30* .0430**		
	VSFA16				10.91* .0018**	4.23* .0450**	7.27* .0095**	
	VSFA64					4.07* .0494**		
Simon Effect	SimCL				16.20* .0002**			
	SimICR		10.09* .0026**	4.23* .0451**				
	SimCRe			5.22* .0263**	6.33* .0150**			
	SimICRe	6.75* .0123**		8.43* .0054**	17.45* .0001**	7.31* .0093**		

* *F value*

** *Pr > F*

(L1) the current status of the lights (on/off), (L2) if the lights were currently in the desired conditions, (S) if the scales were currently in the desired conditions, (Level) if the amounts in tanks were in the desired conditions, (P_fail) the current status of the pumps (i.e. malfunctioning, operating, and not operating), (P_sup) the current task of the pumps (i.e. supplying which tanks), and (P_plans) future plans for the pumps (which pumps to turn on)

Discussion

Higher scores on the Spatial Cuing test were positively correlated with better MAT task performance and MAT situation awareness. Although the task used in this study was a simplified pilot task, it required the ability to quickly gather information, make decisions, and perform the actions. As the pieces of information were evenly spread out on the entire display, with equal importance, the operator needed to simultaneously observe the information received from all areas of the display. Therefore, the ability to quickly shift attention could help improve task performance and situation awareness.

The ability to attend to items presented in rapid series determines the ability to quickly process such information. Surprisingly, higher scores on the Attentional Blink test were negatively correlated with MAT task performance and MAT situation awareness. This unexpected result may occur due to the different characteristics between the Attentional Blink test and the MAT task. When performing the Attentional Blink test, an operator attended to a series of letters and only responded to the presence of the target letters. There was very minimal thinking involved. However, when performing the MAT task, it was necessary to attend to various items at the same time. The operator was required not only to recognize the targets, but also understand their meanings.

Operators with shorter feature search times tended to have better task performance and higher situation awareness. The results from this study agree with Gopher and Kahneman, (1971, in Endsley and Bolstad, 1994). The system-monitoring task required the operator to search the display for the signs of abnormalities, such as the lights and the scale. Therefore, shorter search times were related to better system monitoring performance. Resource management also depends on visual search ability, such as quickly locating the failed pumps, which would enable an operator to more quickly develop strategies to manage the resource.

Surprisingly, operators with longer conjunctive search times tended to have fewer green light misses. It appears that operators who spent more time carefully searching for conjunctive targets tended to be careful about checking the current status of the system. Therefore, operators who were slower searching tended to exhibit fewer misses.

As expected, operators who scored higher on the Stroop Effect test tended to have better task performance and higher situation awareness. Automatic behavior is involved in stimuli identification stage (Coglab, 2000b), so being able to overcome the automation effect could indicate the ability to quickly process the information. Therefore, operators who could more quickly identify the font color also tended to have faster reaction times and more accurate knowledge regarding the lights and scales.

As expected, operators who were quickly and more accurately responding to the targets tended to be have better task performance and higher situation awareness. Location irrelevancy between the target and the response could slow down the response identification process, resulting in poorer MAT task performance and situation awareness.

The results from the regression analysis showed that all of the five attention test scores were found to significantly affect MAT task performance, and the Visual Search and Simon effect test scores affected MAT situation awareness. This result may support the assumption of the probabilistic relationship between task performance and situation awareness. Factors affecting task performance do not necessarily affect situation awareness.

In summary, this study found that limitations of human attention are significantly correlated with task performance and situation awareness. Operators who scored higher on the Spatial Cuing, Stroop Effect, and Simon Effect tests tended to have better task performance and situation awareness when performing a simulated complex task under various system conditions. However, operators who scored higher the Attentional Blink test tended to have poorer task performance and situation awareness. Higher scores on the Visual Search test were positively associated with high situation awareness, but associated with mixed results on task performance.

All five of attention limitation tests can be used to predict MAT task performance, while only the Visual search and the Simon effect tests may be used to predict MAT situation awareness.

Future investigation is required to confirm the current findings, as well as to provide opportunity for further discovery. Possible studies to be conducted include incorporating various forms of information such as audio into the simulated task used in the current study; incorporating different forms of task settings such as automation failure warning lights, levels of severity if failing to meet the goals, or compensation levels into the simulated task used in this study; incorporating different interface designs; including a motor skill battery test in the experiment; considering biography factors such as ages and work experience as individual differences; and repeating this study using a high fidelity simulation.

References

Carretta, T.R., Perry, C.D. Jr., and Ree, M.J. (1996). Prediction of Situational Awareness in F-15 Pilots, *International Journal of Aviation Psychology* *6(1),* 21-41.

Coglab (2000a). Attentional Blink, wysiwyg://8/http://coglab.psych...glab/Labs/AttentionalBlink.html, Date accessed: 12/10/2000.

Coglab (2000b). Stroop Effect, wysiwyg://20/http://coglab.psyc...u/coglab/Labs/StroopEffect.html, Date accessed: 12/10/2000.

Comstock, J.R., Jr., and Arnegard, R.J. (1992). *The Multi-variable Attribute Task battery for Human Operation workload and Strategic Behavior Research.* NASA Technical Memorandum 104174. Hampton VA: National Aeronautics and Space Administration, Langley Research Center.

Endsley, M.R., and Garland, D.J. (2000). Situation Awareness Analysis and Measurement, *LawrenceErlbaum Associates,*New Jersey.

Endsley, M.R. (1988). Design and Evaluation for Situation Awareness Enhancement, *Proceedings of the Human Factors Society 32th Annual Meeting,* **1**, 97-101.

Endsley, M.R. (1995a). Toward the Theory of Situation Awareness in Dynamic Systems, *Human Factors,* **37***(1),* 32-64.

Endsley, M.R., and Bolstad, C.A. (1994). Individual Differences in Pilot Situation Awareness, *The International Journal of Aviation Psychology,* **4***(3)*, 241-264.

Endsley, M.R., and Rogers, M.D. (1998). Distribution of Attention, Situation Awareness, and Workload in a Passive Air Traffic Control task: Implication for Operational Errors and Automation. *Air Traffic Control Quarterly,* **6***(1),* 21-44.

Endsley, M.R., and Smith, R.P., (1996). Attention Distribution and Decision Making in Tactical Air Combat, *Human Factors,* **38***(2),* 232-249.

Fracker, M.L. (1989). Attention Allocation in Situation Awareness, In *Proceedings of the Human Factors Society 33th Annual Meeting,* (pp. 1396-1400). Santa Monica, CA: Human Factors and Ergonomics Society.

Gugerty, L. (1998). Evidence from a Partial Report Task for a Forgetting in Dynamic Spatial Memory, *Human Factors,* **40***(3),* 498-508.

Jones, D.G., Endsley, M.R. (1996). Source of Situation Awareness Errors in Aviation, *Aviation Space Environment Med,* **67**, 507-512.

O'Hare, D. (1997). Cognitive Ability Determinants of Elite Pilot Performance, *Human Factors* **39***(4),* 540-552

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology,* **32**, 3-25.

Raymond, J.E., Sharpiro, K.L., and Aemell, K.M. (1992). Temporary Suspension of Visual Processing in a RSVP Task : An Attentional Blink?. *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 849-860.

Sandom, C. (1999). Situation Awareness Through the Interface: Evaluating System in Safety Critical Control Systems. *IEE Conference Publication*, **463**, 119-124.

Schutte, P.C., Trujillo, A.C. (1996). Flight Crew Task Management in Non-normal Situation, *Proceedings of the Human Factors Society 40th Annual Meeting*, 244-248.

Chapter 6. Task Performance and Situation Awareness: The Role of Memory Limitations

A paper to be submitted to *Human Factors*

Jaruwan Klamklay and Patrick E Patterson

Abstract

This study examined the effects of human memory limitations on situation awareness and performance for operators of a simulated complex task. A simplified pilot task which required operators to monitor system status and manage system fuel resources was used. Eighty participants performed the simulated task under different task pace, duration, and automation levels. Speed and accuracy were used to measure task performance, with the Situation Awareness Global Assessment Technique (SAGAT) used to quantify situation awareness. The subjects also completed four psychological tests to measure memory limitations. The associations between task performance and each of four memory limitations, and between situation awareness and each of the four memory limitations, were quantified using Pearson correlation coefficients. The results showed that operators with high scores on each of the four memory tests tended to exhibit better task performance, and those with higher scores on three specific tests tended to demonstrate higher situation awareness. The regression analysis suggested that the scores from the Memory Span, the Serial Position, and the Absolute Identification tests could predict task performance and situation awareness of operators of a simulated complex system.

Introduction

Situation awareness is needed for operators to perform tasks effectively (Endsley 2000). To reach a goal, the operator must begin by forming situation awareness, making a decision, and then performing necessary actions. These three steps occur continuously and are highly interrelated, therefore the quality of each step depends on the previous one (Endsely,1988; Endsley, 1995).

The importance of situation awareness becomes greater as the complexity and dynamic nature of a system increases. Due to the interconnection and interdependence among system components, a single component error or failure could result in numerous side effects. Additionally, an individual's current situation awareness affects the way the new information is perceived and interpreted (Sandom, 1999 & Endsley, 1988).

Operators of a complex system must continuously keep track of what is transpiring, anticipating potential problems, and preparing to solve any problems encountered. When a problem does occur, operators need to quickly make a decision and perform corrective actions to bring the system back to normal conditions. Incomplete or inaccurate SA will require more time to revise situation awareness, to make a decision, and perform necessary actions. The worst case scenario occurs when operators cannot successfully perform all corrective actions within necessary time limitations, which can lead to disastrous consequences.

Despite its importance, very little is known about SA, and such knowledge is generally based on general intuition (Sarter and Woods, 1991). In addition, few studies have been conducted to confirm the SA construct. More precise understanding of SA is needed for investigating accidents, designing controls and displays to support high SA, designing systematic training programs, or recruiting operators potentially to have high SA.

Among the factors hypothesized to affect situation awareness, memory is considered to be a major cognitive mechanism important to the development of SA (Endsley, 1988, 1995; Endsley and Garland, 2000). All three levels of situation awareness are supported by working memory: perceiving, comprehending, and projecting the future (Adams, Tenny, and Pew, 1995). However, limitations of working memory greatly effect operators' situation awareness. Jones and Endsley (1996) reported that about 8.4% of SA errors occurred due to limitations of memory. Operators reported that they had perceived the information, but it was then forgotten.

Like working memory, long-term memory plays an integral role in situation awareness. The operator use facts regarding the system to build mental models of the system in long-term memory. For example, the description of system purpose, functions, forms, components, and interactions between components. When faced with any situation, this advance knowledge facilitates the perception, the comprehension, and the projection of the future to be processed faster and more accurately (Endsley, 1995). As situation awareness and task performance are related, memory has been hypothesized to effect task performance as well.

Endsley and Bolstad (1994) investigated the effect of memory on pilots' situation awareness. Twenty-five former military fighter pilots were divided into groups of six (some subjects participated on more than one team) and performed air-to-air engagement in a real time simulator. Subject situation awareness was measured using Situation Awareness Global Assessment Technique (SAGAT), with the Immediate/Delayed memory test used to measure short-term memory capacity. A biography survey of subjects' ages, years of flight experience, number of flight hours, and combat experience were included as long-term memory indices. They found no relationships between SA and the Immediate/Delay memory test, but did find a relationship with the four biographical measures.

Carretta, Perry, and Ree (1996) conducted an experiment using 171 USAF pilots who performed an air-to-air mission exercise with their SA collected by using supervisor and peer (pilots rated other pilots) rating methods. The participant's differences in cognitive ability, psychomotor ability, and personality were measured. When holding flying experience constant, they found that general cognitive ability based on working memory, spatial reasoning, and divided attention were significant predictors of SA ($r = 0.7$).

Guterty & Tirre (1996) conducted a study to determine the effect of working memory on situation awareness during a simulated driving task. One hundred and eight US air force trainees performed a driving simulator task and an abstract spatial memory test on a personal computer. The SA measures were correlated with working memory ability ($r = 0.2$, $p < .03$).

O'Hare, (1997) administered the WOMBAT Situation Awareness and Stress Tolerance Test, and the Walter Reed Performance Assessment Battery to a group of 24 males with difference ages, occupations, and computer related experiences. WOMBAT is a complex-multiple task simulation requiring high level of SA. The Walter Reed Performance Assessment Battery includes tests for cognitive abilities--Pattern Recognition, Digit Recall, Six-letter search, and Manikin. Only the Pattern recognition test score was significantly correlated with the WOMBAT performance.

Intuitively, memory should always support operator performance and situation awareness, yet previous studies have found mixed results. It could be hypothesized that the inconsistent results were affected by the methods used to measure memory. Below are details regarding limitations of memory found in previous psychological studies, and the tests developed to demonstrate such limitations.

First, human memory has a limited capacity. Only a limited amount of information can be held in the short-term memory storage. This capacity is called "memory span", a phenomena illustrated by Ellis, Badderley, and Miles (In Baddeley, 1986). A test used to demonstrate this phenomenon is called Memory Span.

Second, there is evidence that humans may remember incidents that in fact never happened, or remember incidents differently from the way they actually happened (Roedigler and MacDermott, 1995; Deese,1959). This phenomenon is called false memory, which can be demonstrated by a test also called False Memory.

Third, the ability to recall items presented in a list is limited by the location of the item in the list. When subjects recall items from a list, in any order, the position of the item in the list has a significant effect on the possibility that an item will be recalled (Kirkpatrick, E.A.,1984). The Serial Position test is used to demonstrate this phenomenon.

Fourth, the ability to identify unidimensional items is limited. When items with one dimension (i.e. weight, tones, or length) are evenly spaced, the human tends to be able to identify the first and the last item more accurately than the items in the middle (Murdock, 1960). The Absolute Identification test is used to demonstrate this phenomenon.

These four memory limitations may directly affect operator performance and situation awareness. Yet, none of the previous studies have actually examined their effects on operator performance or situation awareness.

The first objective of this study was to determine the relationship of these four memory limitations with operator performance and situation awareness. Four hypotheses were investigated: (a) higher scores on the Memory Span test will be positively correlated with better task performance and higher situation awareness, (b) higher scores on the False Memory test will be positively correlated better task performance and higher situation awareness, (c) higher scores on the Serial Position test will be positively correlated better task performance and higher situation awareness, and (d) high scores on the Absolute Identification test be positively correlated better task performance and higher situation awareness.

The second objective was to determine the effect of memory limitations on task performance and situation awareness of operator performing the MAT task under different conditions of automation level, pace, and work duration.

Findings from this study will contribute to a better understanding of the SA construct in complex systems, with additional applications in the improvement of processes for selecting more capable operators of complex systems, in building more efficient and systematic training programs for such systems, and in the design of more effective user interfaces for complex human-information systems.

Methods

Participants

Eighty undergraduate engineering students (25 female and 55 males) ages between 19-25 years from various disciplines were enrolled on a voluntary basis. The study was evaluated and approved by the Iowa State University Human Subject Committee.

Tasks

The operator was required to perform two types of tasks; the Multi-variable Attribute Task battery for Human Operation workload and Strategic Behavior Research (MAT) simulation task, and four psychological tests.

The MAT simulation task

This study used the Multi-variable Attribute Task battery for Human Operation Workload and Strategic Behavior Research (Comstock and Arnegard, 1992) to simulate a simple complex system. Each operator of the system was required to perform two sub-tasks. The first sub-task was system monitoring, in which the operator monitored two lights and four vertical scales for signs of system abnormalities. Under normal conditions, the left light would always be green. However, if a system abnormality occurred, the green light would go out. The operator responded to the disappearance of the green light by pressing the assigned key as soon as possible. The right light was normally off; however, if a system abnormality occurred a red light would turn on in that position. Similarly, once the operator noted the red light was on, he/she pressed the assigned key as soon as possible. The operator would see feedback as the green light would immediately turn back on, and the red light would immediately turn off.

In addition to the lights, the operator also monitored the vertical scales below the lights. Each scale had a yellow pointer that could fluctuate one unit below to one unit above the centerline. If a system fault occurred, the corresponding scale would shift out of the normal range. Regardless of the abnormal shifting direction, the operator could correct the fault by pressing the assigned key. The feedback to a correct response was given by the presence of a yellow bar at the bottom of the scale that indicated an out of range condition, and a return to the center of that scale pointer. If the operator did not notice the fault events (i.e. the lights, the scales), the situation would return to normal condition after a selected time-out period.

The second sub-task was resource management. This sub-system contained six fuel tanks connected by pipelines. The main tanks were tanks A and B, each having 4000-unit capacity. The supply tanks C and D contained a maximum of 2000 units each, and two additional unnamed supply tanks had unlimited capacities. The system used the fuel from tanks A and B, so fuel levels in these two tanks continuously decreased while the system was operating. The operator's task was to maintain the fuel in tanks A and B at their specific levels, as well as trying to keep tanks C and D full. Both numbers underneath tanks A, B, C, and D, and the green levels in the tanks, represented the current amount of fuel in the tanks. To meet the goal, the operator used the eight pumps provided to transfer fuel between tanks. The pumps transferred fuel from one tank to others in the directions indicated by arrows on the display. Keys numbered 1 to 8 were used as toggle switches to turn the corresponding pumps On/Off. No interactions from the operator were required if resource management was in its automatic mode.

The experimenter could set several system parameters including the disappearance of the green light, the appearance of the red light, the out-of-range fluctuation of the scales, time-out periods, resource management automatic mode, pump failures, service time to the failed pumps, pump flow rates, and the fuel consumption rates of tanks A and tank B.

Task performance was measured from the average reaction times to detect and correct system malfunctions (i.e. the scales, the green light, and the red light), proportions of misses (i.e. number of uncorrected malfunctions/ total number of malfunctions), and the average deviation of fuel levels of tanks A, B, C, and D from their target levels.

Each operator's SAGAT score was calculated by comparing the operator's answers to what actually happened. Each query was graded independently. Seven groups of questions presented to the subjects were: (1) the current status of the lights (on/off), (2) if the lights were currently in the desired conditions, (3) if the scales were currently in the desired conditions, (4) if the fuel amounts in the tanks were in the desired ranges, (5) the current status of the pumps (i.e. malfunctioning, operating, and not operating), (6) the current tasks of the pumps (i.e. supplying which tanks), and (7) future plans for the pumps (which pumps to turn on).

Questions were presented as either Yes/No questions, or as multiple answer questions. For Yes/No questions, if the question was answered correctly the score would be one, otherwise zero. For the questions that contain more than one possible answer, the score was determined by the number of correct answers divided by the summation of the number of correct and incorrect answers. Possible answers that were not chosen were considered incorrect answers. The scores for this type of question ranged between zero and one.

Psychological tests

Four psychological tests were administered to measure the limitations of participant memory.

Memory Span. Several series, each containing only one of the following types of stimuli: numbers, letters sounding different, letters sounding the same, short words, and long words, were randomly presented to the subject. After completing each series, the subject recalled the items in the series. The most accurate result from each type of series recalled was recorded with the higher value indicating better use of memory resources.

False Memory. Each subject was presented with several sequences of words. Upon each sequence completion, the subject indicated which words were in the sequence just seen from a pool of words. The average percentages of time a subject chose (1) words that were in the sequence, (2) words that were not in the sequence, and (3) special words very similar to words in the sequence were recorded with higher percentage of (1) and lower percentages of (2) and (3) indicating better use of memory resources.

Serial Position. Each subject was presented with several series of ten letters. In each series, the subject recalled the letters that were in the list in any order. The average percentage of time a subject correctly recalled an item at each position in the sequence was recorded with higher percentage indicating better use of memory resources.

Absolute Identification. Each subject was first presented with a series of seven lines, starting from the shortest line (line 1), to the longest line (line 7). After previewing the seven lines, one of the seven lines was randomly presented to the subject and then repeated several times. Each time the subject identified which line was presented. The average percentages of time a subject could correctly identify line 1 to line 7 were recorded with higher percentages indicating better use of memory resources.

Experimental design

The MAT task contained 2 levels of pace, 2 levels of task duration, and 2 levels of automation, creating 8 possible treatments. Each subject was assigned to perform two different simulation treatments, using balanced incomplete block design. To prevent the participant from guessing the nature of the simulation, each of the two treatments was randomly chosen to perform on different days with at least three days between treatments.

Four psychological tests were administered to each of the subjects in random order. The dependent variables were the operators' MAT task performance and MAT situation awareness. The predictor variables were pace, automation, task duration, and the scores from the four memory tests.

Procedure

The first session began with a 15-minute introduction and practice, allowing the participants to become familiar with the MAT program, the NASA-TLX workload rating scale, and the SAGAT measurement. MAT system components and their fundamental characteristics were explained. The subjects were informed in advance to expect system failure events such as the disappearance of the green light, the appearance of the red light, the deviations of the scales, pump failures, and automation mode failures. Each subject practiced responding to the event faults and controlling the pumps via the computer keyboard. Next, the NASA-TLX self-rating procedure was introduced including the meaning of each sub-scale. Each subject learned how to use the NASA-TLX scale, followed by a practice session. The SAGAT procedure was then introduced to the subjects. The subjects were informed to expect freezes and questions during the experimental session. The subjects were encouraged to answer as many as questions as they could at each pause, and provide their best guess if they did not know the answer. Approximately ten questions were to be expected during each three-minute pause. After the subjects were familiarized with the MAT system, the NASA-TLX, and the SAGAT procedure, they performed a 5-minute practice session. A longer practice session could be performed upon individual request.

After a subject felt comfortable with the task to be performed, the actual situation awareness experiment began. The subject monitored the green light, the red light, the scales, and tried to keep the amount of fuel in tanks A, B, C, and D at the desired levels. Three to six four-minute freezes randomly occurred during the simulation. The subject spent the first minute of the freeze rating their perceived workload using the NASA-TLX rating scale. The three remaining minutes were used to complete SAGAT queries. The subject answered as many questions as possible, guessing if necessary. After four minutes, the simulation resumed. The participant then took a short break before performing the battery of psychological tests.

The second testing session began with a five-minute review and practice. After a warm up period, the participant performed another trial of the MAT battery task, followed by a short break, and then the remaining psychological tests were conducted.

Operator performances were collected as reaction times to the lights and scales, proportions of light and scale misses, and the average deviations of fuel levels from the target levels. Operator situation awareness was measured by the SAGAT method.

Results

The JMP statistical software (JMP 4.0.2) was used to determine the relationships between the factors of interest, as Pearson correlation coefficients. Subject random block effects may affect the Pearson correlation coefficients so only the data from one treatment was randomly chosen to be included in the data analysis.

Table 1. The significant correlations between the memory test scores and task performance ($p < .05$)

Test scores		Task Performance									
		Reaction times			Misses			Deviations			
		S	R	G	MS	MR	MG	D_A	D_B	D_C	D_D
Memory Span	Numbers			-.2221							
	Long Words				-.2265						
False Memory	Lure	.2311									
Serial Position	Position 10										-.2260
Absolute Identification	Line 4								-.2565		
	Line 5				-.2213				-.2628		
	Line 6				-.2202						

S = Scales, R = Red light, G = Green light, MS = Scale misses, MR = Red light misses, MG = Green light misses, and D_I = The deviation of fuel in tank I from its target level

Table 2. The significant correlations between the memory test scores and situation awareness ($p < .05$)

Test scores		Situation Awareness						
		Lights and Scales			Pumps and Fuel Levels			
		L1	L2	S	P_fail	P_sup	Level	P_plans
False Memory	Lure				-.2538			
Serial Position	Position 5					.2282		
	Position 9					.3549		
	Position 10					.2629		
Absolute Identification	Line 4		.3249					

(L1) the current status of the lights (on/off), (L2) if the lights were currently in the desired conditions, (S) if the scales were currently in the desired conditions, (Level) if the amounts in tanks were in the desired conditions, (P_fail) the current status of the pumps (i.e. malfunctioning, operating, and not operating), (P_sup) the current task of the pumps (i.e. supplying which tanks), and (P_plans) future plans for the pumps (which pumps to turn on)

Table 3. A matrix of correlation coefficients between each pair of the memory test scores

	NUM	LD	LS	SW	LW	P1	P2	P3	P4	P5	P6	P7	P8	P9
NUM	1	0.3209	0.221	0.4021	0.2598	0.1227	0.1486	0.1659	0.1626	0.1625	0.2835	0.191	0.1298	0.0691
LD	0.3209	1	0.4172	0.4246	0.3202	0.2788	0.3144	0.349	0.4298	0.3229	0.3655	0.3626	0.3161	0.1677
LS	0.221	0.4172	1	0.4509	0.2895	0.1267	0.1389	0.2239	0.3235	0.2527	0.3191	0.2182	0.288	0.1293
SW	0.4021	0.4246	0.4509	1	0.261	0.2069	0.1149	0.1652	0.2597	0.3188	0.3969	0.3984	0.298	0.1102
LW	0.2598	0.3202	0.2895	0.261	1	0.2504	0.2495	0.3611	0.3068	0.3169	0.2825	0.1803	0.4093	0.1427
P1	0.1227	0.2788	0.1267	0.2069	0.2504	1	0.6664	0.706	0.567	0.5656	0.4853	0.3464	0.2593	0.1639
P2	0.1486	0.3144	0.1389	0.1149	0.2495	0.6664	1	0.7637	0.6233	0.5049	0.5489	0.3305	0.3401	0.2184
P3	0.1659	0.349	0.2239	0.1652	0.3611	0.706	0.7637	1	0.7337	0.6758	0.5649	0.4139	0.4313	0.1649
P4	0.1626	0.4298	0.3235	0.2597	0.3068	0.567	0.6233	0.7337	1	0.6455	0.6074	0.421	0.4377	0.1663
P5	0.1625	0.3229	0.2527	0.3188	0.3169	0.5656	0.5049	0.6758	0.6455	1	0.6504	0.5218	0.5085	0.2344
P6	0.2835	0.3655	0.3191	0.3969	0.2825	0.4853	0.5489	0.5649	0.6074	0.6504	1	0.6922	0.6493	0.4348
P7	0.191	0.3626	0.2182	0.3984	0.1803	0.3464	0.3305	0.4139	0.421	0.5218	0.6922	1	0.7253	0.5124
P8	0.1298	0.3161	0.288	0.298	0.4093	0.2593	0.3401	0.4313	0.4377	0.5085	0.6493	0.7253	1	0.6551
P9	0.0691	0.1677	0.1293	0.1102	0.1427	0.1639	0.2184	0.1649	0.1663	0.2344	0.4348	0.5124	0.6551	1
P10	-0.014	0.0581	0.1144	0.1228	0.1564	0.0633	0.0413	0.0192	0.034	0.1171	0.2528	0.3474	0.5887	0.637
Line1	0.0533	0.1952	0.235	0.0472	0.0795	-0.035	0.0993	0.0735	0.0851	-0.02	0.0682	-0.02	0.0097	0.0747
Line2	0.0172	0.1915	0.197	-0.003	0.0753	-0.089	-0.025	-0.009	-0.163	-0.1	-0.15	-0.126	-0.084	0.004
Line3	-0.068	0.2605	0.1495	0.0046	0.1988	-0.006	0.0176	0.0274	0.0747	0.0141	0.0428	0.025	0.0684	-0.006
Line4	0.1844	0.376	0.3899	0.1694	0.3952	0.1434	0.2934	0.3344	0.3398	0.3323	0.1893	0.0195	0.1379	0.0373
Line5	0.1329	0.1864	0.1412	0.1784	0.1288	0.1102	-0.004	0.1036	0.066	0.0528	-0.019	0.0205	0.0821	0.1291
Line6	0.0011	0.1423	0.1428	0.0805	-0.009	0.1797	0.0769	0.2481	0.1726	0.1239	0.0726	0.0429	0.1278	0.2986
Line7	0.1318	0.1235	0.0705	0.0595	-0.001	0.19	0.1878	0.1164	0.1194	0.1453	0.1001	0.1203	0.1219	0.1602
Wlisted	0.1693	0.1739	0.1406	0.1907	0.3472	0.2035	0.1409	0.1942	0.2499	0.2028	0.1861	0.0771	0.1181	-0.005
WnListed	-0.076	-0.085	-0.145	0.0224	0.0126	-0.189	-0.131	-0.128	-0.068	0.0139	-0.168	-0.083	-0.113	-0.268
Lure	0.1515	-0.071	0.0671	0.0332	0.0618	0.0124	-0.055	-0.03	-0.058	-0.056	-0.044	0.0375	-0.082	-0.185

Table 3 (continued). A matrix of correlation coefficients between each pair of the memory test scores

	P10	Line1	Line2	Line3	Line4	Line5	Line6	Line7	Wlisted	WnListed	Lure
NUM	-0.014	0.0533	0.0172	-0.0679	0.1844	0.1329	0.0011	0.1318	0.1693	-0.0761	0.1515
LD	0.0581	0.1952	0.1915	0.2605	0.376	0.1864	0.1423	0.1235	0.1739	-0.085	-0.0708
LS	0.1144	0.235	0.197	0.1495	0.3899	0.1412	0.1428	0.0705	0.1406	-0.1452	0.0671
SW	0.1228	0.0472	-0.0034	0.0046	0.1694	0.1784	0.0805	0.0595	0.1907	0.0224	0.0332
LW	0.1564	0.0795	0.0753	0.1988	0.3952	0.1288	-0.0089	-0.001	0.3472	0.0126	0.0618
P1	0.0633	-0.0345	-0.089	-0.0059	0.1434	0.1102	0.1797	0.19	0.2035	-0.1885	0.0124
P2	0.0413	0.0993	-0.0246	0.0176	0.2934	-0.0036	0.0769	0.1878	0.1409	-0.1313	-0.0547
P3	0.0192	0.0735	-0.0085	0.0274	0.3344	0.1036	0.2481	0.1164	0.1942	-0.1283	-0.03
P4	0.034	0.0851	-0.1633	0.0747	0.3398	0.066	0.1726	0.1194	0.2499	-0.0684	-0.058
P5	0.1171	-0.0195	-0.1	0.0141	0.3323	0.0528	0.1239	0.1453	0.2028	0.0139	-0.0557
P6	0.2528	0.0682	-0.1503	0.0428	0.1893	-0.0186	0.0726	0.1001	0.1861	-0.1681	-0.0439
P7	0.3474	-0.0195	-0.1262	0.025	0.0195	0.0205	0.0429	0.1203	0.0771	-0.0829	0.0375
P8	0.5887	0.0097	-0.0841	0.0684	0.1379	0.0821	0.1278	0.1219	0.1181	-0.113	-0.0819
P9	0.637	0.0747	0.004	-0.0064	0.0373	0.1291	0.2986	0.1602	-0.0046	-0.2677	-0.1852
P10	1	0.0768	0.0937	0.1796	-0.0488	0.1872	0.1652	0.1359	0.1682	-0.1144	-0.0324
Line1	0.0768	1	0.2143	0.2672	0.1928	0.1474	0.2168	0.1369	-0.0706	-0.1286	-0.2415
Line2	0.0937	0.2143	1	0.4693	0.3207	0.3755	0.1391	0.1661	0.016	-0.077	-0.1582
Line3	0.1796	0.2672	0.4693	1	0.2187	0.3157	0.0995	0.1698	0.1872	-0.1247	-0.0043
Line4	-0.049	0.1928	0.3207	0.2187	1	0.4148	0.2406	0.2765	0.2001	-0.0327	-0.1291
Line5	0.1872	0.1474	0.3755	0.3157	0.4148	1	0.4044	0.2083	0.2681	-0.3324	-0.0439
Line6	0.1652	0.2168	0.1391	0.0995	0.2406	0.4044	1	0.3231	0.2249	-0.3431	-0.0554
Line7	0.1359	0.1369	0.1661	0.1698	0.2765	0.2083	0.3231	1	0.0773	-0.2122	0.218
Wlisted	0.1682	-0.0706	0.016	0.1872	0.2001	0.2681	0.2249	0.0773	1	0.0471	0.3521
WnListed	-0.114	-0.1286	-0.077	-0.1247	-0.0327	-0.3324	-0.3431	-0.2122	0.0471	1	0.0503
Lure	-0.032	-0.2415	-0.1582	-0.0043	-0.1291	-0.0439	-0.0554	0.218	0.3521	0.0503	1

Table 4. The significant effects of memory test scores on task performance ($p < .05$)

Test scores		**Task Performance**									
		Reaction times			**Misses**			**Deviations**			
		S	R	G	MS	MR	MG	D_A	D_B	D_C	D_D
Memory Span	Number								4.11* .0473**		
Serial Position	Position 2						10.00* .0026**				
	Position 5						5.91* .0184**				
Absolute Identification	Line 3						4.27* .0433**				

* *F value*

** *Pr > F*

S = Scales, R = Red light, G = Green light, MS = Scale misses, MR = Red light misses, MG = Green light misses, and

D_I *= The deviation of fuel in tank I from its target level*

Table 5. The significant effects of memory test scores on situation awareness ($p < .05$)

Test scores		Situation Awareness						
		Lights and Scales			Pumps and Fuel Levels			
		L1	L2	S	P_fail	P_sup	Level	P_plans
Memory Span	Short Word			4.79* .0330**				
	Letters sound different				4.45* .0398**			
	Letters sound the same			4.27* .0436**				
Serial Position	Position 4	5.91* .0186**						
	Position 8	4.80* .0328**						8.10* .0062**
	Position 9			5.33* .0248**	4.68* .0352**	5.30* .0251**		
	Position 10	4.96* .0301**						
Absolute Identification	Line 5				4.36* .0420**			

** F value*

*** Pr > F*

(L1) the current status of the lights (on/off), (L2) if the lights were currently in the desired conditions, (S) if the scales were currently in the desired conditions, (Level) if the amounts in tanks were in the desired conditions, (P_fail) the current status of the pumps (i.e. malfunctioning, operating, and not operating), (P_sup) the current task of the pumps (i.e. supplying which tanks), and (P_plans) future plans for the pumps (which pumps to turn on).

Relationships between memory limitations and task performance

The results (Table 1) show that higher scores on the Memory Span test were negatively associated with the reaction time to the green light and the proportion of scales misses. Operators with larger memory capacity on Long Words and Numbers tended to have shorter reaction time, and had fewer misses. Higher scores on the False Memory test were positively associated with the reaction time to the scales. Operators who reported fewer false memories tended to have shorter reaction time.

Higher scores on the Serial Position were negatively associated with the deviation of fuel level in tank D. Operators who could better recall the list items tended to be better in maintaining the fuel of tank D close to the desired level. Higher scores on the Absolute Identification test were negatively correlated with proportions of the scales misses and the deviation of fuel level in tank B. Operators who could better identify the items in the list tended to have fewer scales misses and better in maintaining the fuel of tank B close to the desired level. In summary, individual who scored higher on each of the memory test performed the MAT task more efficiently.

Relationships between memory limitations and situation awareness

The results (Table 2) show that higher scores on the False Memory test indicated better situation awareness. Operators who less frequently recalled the special lure tended to have more accurate knowledge regarding the current status of the pumps. Higher scores on the Serial Position test indicated better situation awareness. Operators who could more accurately recall the items in the list tended to have more accurate knowledge regarding pumps. Higher scores on the Absolute Identification test indicated better situation awareness. Operator who could more accurately identify one-dimensional items in the list tended to have more accurate knowledge regarding the lights. There was no significant relationship between Memory span test and situation awareness.

Effects of memory limitations on task performance and situation awareness

A multiple regression mixed model was created for each of the response variables, using the SAS statistical program. Measurements from each of the memory tests and treatment factors were included as predictor variables, with subjects as random effects. All 160 data points were included in the analysis. Table 3 shows a matrix of correlation coefficients between each pair of the memory test scores.

Effects of memory limitations on task performance. The results (Table 4) show that the Memory Span, Serial Position and Absolute Identification test scores significantly affected MAT task performance.

Effects of memory limitations on SA. The results (Table 5) show that performance of the Memory Span, Serial Position, and Absolute Identification tests significantly affected MAT situation awareness.

Discussion

Overall, operators who scored higher on the memory limitation tests tended to have better task performance, and higher situation awareness. All four types of memory limitations studied correlated with task performance, and three out of four correlated with situation awareness.

As expected, (1) operators with higher scores on the memory span test tended to exhibit better performance, (2) the operators with higher scores on the False Memory test tended to exhibit better performance, (3) operators with higher scores on the Serial Position test tended to exhibit better performance, and (4) operators with higher scores on the Absolute Identification test tended to exhibit better performance. Better use of memory resource associated with being more efficient in performing MAT task.

Higher scores on the memory tests also indicated better situation awareness. Individuals with larger memory spans could be more efficient in holding the information regarding the status of the system components. Less false memory would increase accuracy when recalling the gathered information. In addition, the ability to recall items in the list and absolutely identify items would help when referring to such data. Each of these abilities could therefore promote better decision making, and aid in performing the necessary actions, resulting in faster response times, more accurate responses, and the ability to maintain fuel at the desired levels.

As expected, higher scores on the False Memory, Serial Position, and Absolute Identification tests indicated better task performance and higher situation awareness. Surprisingly, higher scores on the Memory Span test only supported better MAT task performance, not MAT task situation awareness. The dissociation between memory span and situation awareness was similar to that found in previous studies. O'Hare (1997) found similar results that the Pattern Recognition test correlated with WOMBAT performance while the Digital Recall and the Manikin tests were not. The Pattern Recognition test could be viewed as a mix of the False Memory, the Absolute Identification, and the Serial Position tests while the Digit Recall test as the False Memory tests. Endsley & Bolstad (1994) also found that results from the Immediate/Delay memory test were not correlated with SA.

Hypotheses can be made regarding the dissociation between memory span and SA. First, the nature of the memory span test was different from the MAT task. The test required the operator to stare at the item appearing on a screen one at time. On the contrary, the MAT task used in this study was highly dynamic with, for example, up to four system abnormalities occurring during a one-minute period. Rapidly updating memory may cause the new information to interfere with the old information (Salvendy, 1997, page 106). Moray (1980 in Salvendy, 1997 page 106) found that the memory span of a dynamic memory task was less than five chunks. With such a fast pace, MAT operators may not have time to remember all the system aspects, and therefore their answers for the SAGAT were primarily guesswork.

Second, the atmosphere of the experiment in this study didn't give a real feeling of the possible disastrous consequences of failing to complete a complex task, which could alter the study results.

Lastly, the method used to quantify SA may affect the relationship between memory limitations and SA. Pausing the experimental simulation to ask questions may be considered as noise in the data stream which may impede operator situation awareness (Green, Odom, and Yates, 2000). Sarter and Woods (1995) believed that SAGAT is an intrusive method as queries give clues to the operator of what to attend next.

Another interesting point is that the components of task performance and of situation awareness that correlated with memory limitations. Memory limitations correlated with all aspects of task performance (i.e. both the monitoring performance and the resource management performance). However, memory limitations correlated with mostly situation awareness regarding the resource management task. This may suggest that memory limitations may only correlate with situation awareness of tasks requiring higher cognitive ability rather than tasks requiring simple reaction.

The results from the regression analysis showed that only the scores from the Memory Span, Serial Position and Absolute Identification tests were found to significantly affect task performance, and situation awareness. This result supports the link between task performance and situation awareness that better task performance may lead to higher situation awareness.

In conclusion, the limitations of human memory were significantly correlated with situation awareness and task performance. Individuals who scored higher on each of the four memory limitation tests tended to exhibit better task performance when performing a simulated complex task under varying system conditions. Individuals with higher scores on the False Memory, Serial Position, and Absolute Identification tests tended to have higher situation awareness. The Memory Span, Serial Position, and Absolute Identification test scores could be used to predict the MAT performance and situation awareness.

Further investigation is required to confirm the current findings, as well as to provide opportunity for further discovery. Possible studies to be conducted include:

- Incorporating audio into the simulated task used in the current study
- Incorporating different forms of task settings such as an automation failure warning light
- Incorporating severity levels if operators fail to meet goal, or compensation levels into the simulated task used in this study
- Incorporating different interface designs
- Include a motor skill battery test in the experiment
- Expand the effect of individual differences on situation awareness and task performance of operators of a high fidelity simulation.

References

Adams, M.J., Tenny, Y.J., Pew, R.W. (1995). Situation Awareness and the Cognitive Management of Complex Systems. *Human Factors,* **37***(1),* 85-104.

Baddeley, Alan. (1986). *Working Memory.* Oxford University Press New York USA

Carretta, T.R., Perry, C.D. Jr., and Ree, M.J. (1996). Prediction of Situational Awareness in F-15 Pilots, *International Journal of Aviation Psychology* **6***(1),* 21-41.

Comstock, J.R., Jr., and Arnegard, R.J. (1992). *The Multi-variable Attribute Task battery for Human Operation workload and Strategic Behavior Research.* NASA Technical Memorandum 104174. Hampton VA: National Aeronautics and Space Administration, Langley Research Center.

Deese, J. (1959). On the Prediction of Occurrence of Particular Verbal Intrusions in Immediate Recall. *Journal of Experimental Psychology,* **58**, 17-22.

Endsley, M.R., and Garland, D.J. (2000). Situation Awareness Analysis and Measurement, *LawrenceErlbaum Associates,* New Jersey.

Endsley, M.R. (1988). Design and Evaluation for Situation Awareness Enhancement, *Proceedings of the Human Factors Society 32th Annual Meeting,* **1**, 97-101.

Endsley, M.R. (1995). Toward the Theory of Situation Awareness in Dynamic Systems, *Human Factors,* **37***(1),* 32-64.

Endsley, M.R., and Bolstad, C.A. (1994). Individual Differences in Pilot Situation Awareness, *The International Journal of Aviation Psychology,* **4***(3)*, 241-264.

Green, M., Odom, J.V., and Yates, J.T. (2000). Measuring Situation Awareness with the "Ideal Observer", *http://www.ergogero.com/sitaw/sitaware.html,* Date accessed: 11/27/2000.

Gugerty, L.J., and Tirre, W.C. (1996). Situation Awareness: A Validation Study and Investigation of Individual Differences. In *Proceedings of the Human Factors Society 40th Annual Meeting,* (pp. 564-568). Santa Monica, CA: Human Factors and Ergonomics Society.

Jones, D.G., Endsley, M.R. (1996). Source of Situation Awareness Errors in Aviation, *Aviation Space Environment Med,* **67**, 507-512.

Murdock, B.B. (1960). The Distinctiveness of Stimuli. *Psychological Review,* **67**, 16-31.

Roediger, H. L. III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 803-814.

O'Hare, D. (1997). Cognitive Ability Determinants of Elite Pilot Performance, *Human Factors* **39***(4)*, 540-552.

Salvendy, G. (1997). *Handbook of Human Factors and Ergonomics* 2nd Edition. John Wiley&Sons. New York

Sandom, C. (1999). Situation Awareness Through the Interface: Evaluating System in Safety Critical Control Systems. *IEE Conference Publication*, **463**, 119-124.

Sarter, N.B. (1990). Situation Awareness: A critical but Ill-Defined Phenomenon. In *Proceedings of the Human Factors Society 34th Annual Meeting*, (p23). Santa Monica, CA: Human Factors and Ergonomics Society.

Sarter, N.B., and Woods, D.D. (1991). Situation Awareness: A Critical but ill-Defined Phenomenon. *International Journal of Aviation Psychology*, **1**, 45-57.

Sarter, N.B., and 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.

Chapter 7. The Effect of Individual Differences on Association of Workload, Operator Performance, and Situation Awareness

A paper to be submitted to *Human Factors*

Jaruwan Klamklay and Patrick E Patterson

Abstract

The objective of this study was to identify the relationships among workload, task performance, and situation awareness, with and without taking individual differences into consideration. Eighty participants performed a simplified pilot task, in which response accuracy and latency were used to measure operator performance, and the NASA-TLX to measure workload. The Situation Awareness Global Assessment Technique (SAGAT) and a bi-polar subjective rating scale were used as measures of situation awareness. In addition, nine psychological tests were used to measure individual differences of attention and memory. The association between workload, performance, and situation awareness of the entire participant pool and of participants with similar attention and memory capacities were identified using Pearson correlation coefficients. The results suggest that individual differences play a significant role in defining the relationships amongst workload, performance, and situation awareness. They resulted in not only stronger correlation coefficients amongst workload, task performance, and situation awareness, but also produced different sets of correlated components.

Introduction

Operator workload, performance, and situation awareness are the principle approaches used in design evaluation (Selcon, Taylor, and Koritsas,1991; Endsley, Selcon, Hardiman, and Croft, 1998). If these three approaches were measured during system evaluation, findings from all three approaches would be needed to ensure an optimal design. A good design should generate an acceptable amount of workload that will allow operators to maximally acquire and maintain situation awareness regarding crucial system elements and goals, and will allow operators to reach the desired level of performance.

The measurement of workload, performance and situation awareness requires an investment of time and resources. Subjective rating scales such as NASA-TLX and SWAT are generally used to measure workload because they are inexpensive, sensitive to change of workload levels, easy to administer, are not intrusive, and can be applied to wide variety of tasks. However, unlike measuring workload, each system may require its own specific tools to measure operator performance and situation awareness. Knowledge of the relationships between workload, performance, and situation awareness therefore could reduce the necessity of considering all three approaches during the system evaluation.

Raby and Wickens (1994) determined the affect of task generated workload on aviation task performance, and found that pilot performance measured as tracking error, glideslop error, localiser error, and air speed error decreased as workload increased.

Linkage between workload and task performance also has been found in system domains besides aviation. Hancock (1996) studied the effect of perceived workload on tracking task performance of students and university staff. Operators exhibited better performance (less tracking error) under conditions that they rated as low workload. Hancock also found that the perceived workload gradually decreased across the continual trials, while task performance gradually increased.

Becker et al (1991) studied the effects of feedback on perceived workload during vigilance performance. Feedback about good performance (Hit) resulted in a lower perceived workload, while feedback about bad performance (i.e. misses) resulted in higher perceived workload.

A connection between workload and situation awareness has also been found. Selcon et al (1991) presented three videotaped computer-graphic air combat flight simulation sequences to 12 Royal Air Force (RAF) pilots. Each subject pretended he was the pilot of the aircraft in the video, using the NASA-TLX to quantify perceived workload and the SART subjective method to quantify situation awareness after each presentation was over. They found situation awareness was significantly related to workload.

It is perhaps more reliable to use an objective measure of SA. Endsley (2000) used the Situation Awareness Global Assessment Technique (SAGAT) as a measure of situation awareness, and the NASA-TLX as a measure of perceived workload. Ten active controllers performed four conditions of free flight levels, each under two air traffic scenarios. The results showed that controllers tended to have higher situation awareness when performing a low workload free flight condition.

Heers and Casper (1998) demonstrated a connection between subjective workload and situation awareness of helicopter pilots. Eight active-duty helicopter pilots flew different Rotocraft Pilot's Associated Mission Equipment Packages. The NASA-TLX subjective workload and single-scale situation awareness rating methods were used to measure workload and situation awareness respectively. Scenarios that were rated lower in overall workload tended to be rated higher in situation awareness.

Endsley (1993), in disagreement with other studies, did not find the associations between perceived workload on the situation awareness in a study involving six former military pilots. Each pilot performed 25 trials of air-to-air engagements in a real-time simulator facility. The SWAT was used to measure workload and the SAGAT was used to measure situation awareness. The results showed no significant relationship between workload and situation awareness.

Hallbert (1997) also found a dissociation between workload and situation awareness. Hallbert conducted a study to test and evaluate a realistic control room, using eight licensed nuclear power plant operators. The subjects worked as a team, performing five scenarios over a two to three day period. The results showed that during the workload transition periods (i.e. from normal to abnormal conditions), the operators' situation awareness significantly decreased. After a period of time, the SA recovered to about 50%, while the workload remained high.

Although many studies have been conducted to date, the relationships among perceived workload, performance, and situation awareness are confounded in those studies. Therefore, requiring more targeted studies to verify such relationships. The first objective of this study was to identify the relationships between workload and performance, and between workload and situation awareness for system operators. Better task performance and high situation awareness were expected when performing a task perceived generating low workload, and vice versa.

The inconsistency in relationships found in previous studies could come from factors such as small groups of participants, different task domains used, different measurement methods, or individual differences. Nygren, Schnipke, and Reid (1998) found a significant association between perceived workload (measured with SWAT) and task performance. However, such association was affected by how individuals perceived the importance of a workload dimension. SWAT measures workload in three dimensions; time, effort, and stress. Each of 124 students performed three scenarios of tasks in a virtual office building. In each scenario, one of the workload dimensions was manipulated to be more important than the other two dimensions. An association between workload and task performance existed for individuals who perceived a workload dimension as low importance when in reality that particular dimension was manipulated to be more important than other workload dimensions.

The results from Nygren et al (1998) pointed out that individual differences played a significant role in defining the relationship between workload and situation awareness. However, very few studies have determined the effect of individual differences on relationships among workload, performance, and situation awareness. Therefore, more targeted studies are needed to confirm such effects.

The second objective was to examine the relationship between workload and performance, and workload and situation awareness when taking the limitations of attention and memory into consideration.

Attention and memory are critical elements of human information processing, a process closely related to operator situation awareness and performance. Therefore, this study investigated individual differences in attention and memory limitations. Five psychological tests (Attentional Blink, Spatial Cuing, Visual Search, Stroop Effect, and

Simon Effect) were used to measure attention limitations, and four psychological tests (Memory Span, False Memory, Serial Position, and Absolute Identification) were used to measure memory limitations. Speed and/or accuracy were used to score each test. Higher speed and/or accuracy resulting from each test indicated that an individual had better attention or memory usage, depending on the objective of such test.

Four groups of individuals with extreme scores on the attention tests and extreme scores on the memory tests were considered. Each of the four target groups consisted of individuals who had (1) extremely high scores on the attention tests, (2) extremely low scores on the attention tests, (3) extremely high scores on the memory tests, and (4) extremely low scores on the memory tests. The relationships amongst workload, performance, and situation awareness obtained from these four groups of individuals were expected to be different from that obtained from the entire participant pool. An association occurring in any one of the four groups might not always occur in the entire participant pool.

Workload was measured by the NASA-TLX subjective rating method. Operator performance was the compounded measurement of response latencies and error rates. Situation awareness was obtained from two measurements--SAGAT, and a bi-polar subjective rating scale.

A firm knowledge of the relationship between workload, performance, and situation awareness could reduce the necessity of considering all three of these approaches during system evaluation. It could be applied to system evaluation strategies to decrease the cost and time required yet keep the reliability of the system evaluation at an acceptable level.

Methods

Participants

Eighty undergraduate engineering students (25 female and 55 males) ages between 19-25 years from various disciplines were enrolled on a voluntary basis. The study was evaluated and approved by the Iowa State University Human Subject Committee.

Tasks

Subjects were required to perform two types of tasks; the Multi-variable Attribute Task battery for Human Operation workload and Strategic Behavior Research (MAT) simulation task, and nine psychological tests.

The MAT simulation task

This study used the Multi-variable Attribute Task battery for Human Operation Workload and Strategic Behavior Research (Comstock and Arnegard, 1992) to simulate a simple complex system. Each operator of the system was required to perform two sub-tasks. The first sub-task was system monitoring, in which the operator monitored two lights and four vertical scales for signs of system abnormalities. Under normal conditions, the left light would always be green. However, if a system abnormality occurred, the green light would go out. The operator responded to the disappearance of the green light by pressing the assigned key as soon as possible. The right light was normally off; however, if a system abnormality occurred a red light would turn on in that position. Similarly, once the operator noted the red light was on, he/she pressed the assigned key as soon as possible. The operator would see feedback as the green light would immediately turn back on, and the red light would immediately turn off.

In addition to the lights, the operator also monitored the vertical scales below the lights. Each scale had a yellow pointer that could fluctuate one unit below to one unit above the centerline. If a system fault occurred, the corresponding scale would shift out of the normal range. Regardless of the abnormal shifting direction, the operator could correct the fault by pressing the assigned key. The feedback to a correct response was given by the presence of a yellow bar at the bottom of the scale that indicated an out of range condition, and a return to the center of that scale pointer. If the operator did not notice the fault events (i.e. the lights, the scales), the situation would return to normal condition after a selected time-out period.

The second sub-task was resource management. This sub-system contained six fuel tanks connected by pipelines. The main tanks were tanks A and B, each having 4000-unit capacity. The supply tanks C and D contained a maximum of 2000 units each, and two additional unnamed supply tanks had unlimited capacities. The system used the fuel from tanks A and B, so fuel levels in these two tanks continuously decreased while the system was operating. The operator's task was to maintain the fuel in tanks A and B at their specific levels, as well as trying to keep tanks C and D full. Both numbers underneath tanks A, B, C, and D, and the green levels in the tanks, represented the current amount of fuel in the tanks. To meet the goal, the operator used the eight pumps provided to transfer fuel between tanks. The pumps transferred fuel from one tank to others in the directions indicated by arrows on the display. Keys numbered 1 to 8 were used as toggle switches to turn the corresponding pumps On/Off. No interactions from the operator were required if resource management was in its automatic mode.

The experimenter could set several system parameters including the disappearance of the green light, the appearance of the red light, the out-of-range fluctuation of the scales, time-out periods, resource management automatic mode, pump failures, service time to the failed pumps, pump flow rates, and the fuel consumption rates of tanks A and tank B.

Task performance was measured from the average reaction times to detect and correct system malfunctions (i.e. the scales, the green light, and the red light), proportions of misses (i.e. number of uncorrected malfunctions/ total number of malfunctions), and the average deviation of fuel levels of tanks A, B, C, and D from their target levels.

Each operator's SAGAT score was calculated by comparing the operator's answers to what actually happened. Each query was graded independently. Seven groups of questions presented to the subjects were: (1) the current status of the lights (on/off), (2) if the lights were currently in the desired conditions, (3) if the scales were currently in the desired conditions, (4) if the fuel amounts in the tanks were in the desired ranges, (5) the current status of the pumps (i.e. malfunctioning, operating, and not operating), (6) the current tasks of the pumps (i.e. supplying which tanks), and (7) future plans for the pumps (which pumps to turn on).

Questions were presented as either Yes/No questions, or as multiple answer questions. For Yes/No questions, if the question was answered correctly the score would be one, otherwise zero. For the questions that contain more than one possible answer, the score was determined by the number of correct answers divided by the summation of the number of correct and incorrect answers. Possible answers that were not chosen were considered incorrect answers. The scores for this type of question ranged between zero and one.

For subjective rating of SA, subjects were asked to rate their own situation awareness on a scale of 1 to 10, where 1 was the least and 10 was the most, as to how much situation awareness that they think they had.

Psychological tests

Five classic psychological tests were administered to each participant to measure attention limitations.

Attentional Blink. Each subject watched several series of ten letters, in which the newest letter overwrote the previous letter. Each series contained the first target letter, and/or the second target letter. The separation between the presence of the first target and the second target may vary between series. Each subject reported the presence of both target letters. The average percentage of time a subject could correctly identify the second targets was recorded, with greater values indicating better use of attention resources.

Spatial Cuing. Each subject was required to respond to the presence of targets as quickly as possible under three conditions: (1) when a cue correctly identified a location of a target, (2) when a cue incorrectly identified a location of a target, and (3) when there was no cue provided. The average reaction times under these three conditions were recorded with a faster reaction time indicating better use of attention resources.

Visual Search. Each subject performed two types of visual search tasks. The first type was a conjunctive search, in which the targets shared at least one characteristic (i.e. color, or shape) with the background. The second type was a feature search, in which the targets were completely different form the background and, therefore, would "pop-out" of the display. For the conjunctive search, each subject determined whether a green circle was presented among green squares and blue circles. For the feature search, each subject determined whether a green circle was presented among blue squares. The non-target items, such as green squares and blue circles, were distracters. Average reaction times were recorded with faster visual search times indicating better use of attention resources.

Stroop Effect. Each subject was required to identify the font colors of the words "BLUE", "RED", and "GREEN". The words could be written in the font colors blue, red, or green. The average reaction times to identify the font color when the font color and the word name were the same, and when they were different, were recorded with a faster reaction time indicating better use of attention resources.

Simon Effect. Each subject was required to respond to the presence of the target as quickly possible under four different conditions: (1) when the target was presented on the right, and the response key was located on the right, (2) when the target was presented on the left, and the response key was located on the left, (3) when the target was presented on the right, and the response key was located on the left, and (4) when the target was presented on the left, and the response key was located on the right. Average reaction times and average proportion of errors from conditions 1 to 4 were recorded with faster reaction times, and fewer errors indicating better use of attention resources.

Four classic psychological tests were administered to the participants to measure the memory limitations.

Memory Span. Several series, each containing only one of the following types of stimuli: numbers, letters sounding different, letters sounding the same, short words, and long words, were randomly presented to the subject. After completing each series, the subject recalled the items in the series. The most accurate result from each type of series recalled was recorded with the higher value indicating better use of memory resources.

False Memory. Each subject was presented with several sequences of words. Upon each sequence completion, the subject indicated which words were in the sequence just seen from a pool of words. The average percentages of time a subject chose (1) words that were in the sequence, (2) words that were not in the sequence, and (3) special words very similar to words in the sequence were recorded with higher percentage of (1) and lower percentages of (2) and (3) indicating better use of memory resources.

Serial Position. Each subject was presented with several series of ten letters. In each series, the subject recalled the letters that were in the list in any order. The average percentage of time a subject correctly recalled an item at each position in the sequence was recorded with higher percentage indicating better use of memory resources.

Absolute Identification. Each subject was first presented with a series of seven lines, starting from the shortest line (line 1), to the longest line (line 7). After previewing the seven lines, one of the seven lines was randomly presented to the subject and then repeated several times. Each time the subject identified which line was presented. The average percentages of time a subject could correctly identify line 1 to line 7 were recorded with higher percentages indicating better use of memory resources.

Experimental design

The MAT task contained 2 levels of pace, 2 levels of task duration, and 2 levels of automation, creating 8 possible treatments. Each subject was assigned to perform two different simulation treatments, using balanced incomplete block design. To prevent the participant from guessing the nature of the simulation, each of the two treatments was randomly chosen to perform on different days with at least three days between treatments.

Nine psychological tests were administered to each of the subjects in random order. The variables of interest were the operators' MAT task performance, MAT situation awareness, and the scores from the nine psychological tests.

Procedure

The first session began with a 15-minute introduction and practice, allowing the participants to become familiar with the MAT program, the NASA-TLX workload rating scale, a bi-polar SA subjective rating, and the SAGAT measurement. MAT system components and their fundamental characteristics were explained. The subjects were informed in advance to expect system failure events such as the disappearance of the green light, the appearance of the red light, the deviations of the scales, pump failures, and automation mode failures. Each subject practiced responding to the event faults and controlling the pumps via the computer keyboard. Next, the NASA-TLX self-rating procedure was introduced including the meaning of each sub-scale. Each subject learned how to use the NASA-TLX scale, followed by a practice session. The SAGAT, and a bi-polar SA subjective rating procedures were then introduced to the subjects. The subjects were informed to expect freezes and questions during the experimental session. The subjects were encouraged to answer as many as questions as they could at each pause, and provide their best guess if they did not know the answer. Approximately ten questions were to be expected during each three-minute pause. After the subjects were familiarized with the MAT system, the NASA-TLX, and the SAGAT procedure, they performed a 5-minute practice session. A longer practice session could be performed upon individual request.

After a subject felt comfortable with the task to be performed, the actual situation awareness experiment began. The subject monitored the green light, the red light, the scales, and tried to keep the amount of fuel in tanks A, B, C, and D at the desired levels. Three to six four-minute freezes randomly occurred during the simulation. Subjects spent the first minute of the freeze rating their perceived workload using the NASA-TLX rating scale. The three remaining minutes were used to complete SAGAT queries. Participants answered as many as questions as possible, guessing if necessary. After four minutes, the simulation resumed. Each participant then took a short break before performing the battery of psychological tests.

The second testing session began with a five-minute review and practice. After a warm up period, the participant performed another trial of the MAT task, followed by a short break, and then the remaining psychological tests were conducted.

Operator perceived workload was measured using the NASA-TLX subjective rating scale. The operator performances were collected as reaction times to the lights and scales, proportions of light and scale misses, and the average deviation of fuel levels from the target levels. Operator situation awareness was measured by a single bi-polar subjective rating scale, and with the SAGAT methods.

Results

The JMP statistical software (JMP 4.0.2) was used to determine the relationships between the factors of interest, as Pearson correlation coefficients. Subject random block effects may affect the Pearson correlation coefficients so only the data from one treatment was randomly chosen to be included in the data analysis.

Table 1 shows the Pearson correlation coefficients between: (1) task performance and overall perceived workload (TLX), (2) SA and overall perceived workload (TLX), and (3) self rated SA and overall perceived workload (TLX), using the JMP statistical software (JMP 4.0.2). The 5% significance level was used throughout the entire data analysis.

Results for the entire participant pool

Correlation between task performance and workload

None of the task performance elements were significantly correlated with the overall perceived workload. Operator performance had no association with overall perceived workload.

Correlation between situation awareness and workload

None of the SAGAT elements were significantly correlated with the overall perceived workload. Operator knowledge about system aspects had no association with overall perceived workload.

Correlation between rated situation awareness and workload

Self rated SA was significantly correlated with overall perceived workload ($r = -.3445$ $p < .004$). Operators with a higher self rated SA tended to indicate that they had a lower overall perceived workload.

Results for individuals with extreme attention and memory scores

Individuals with extreme attention scores

A multiple regression analysis was conducted to identify which attention test scores affect overall perceived workload. The regression showed that the "average reaction time to respond to the target when the target was presented on the right, while the response key was presented on the left" was significantly affected overall perceived workload. Shorter average reaction time represented higher attention test scores. Two extreme groups of individuals were obtained, a group of individuals with better attention resources coming from the highest

20 attention test scores, and another group of individuals with poorer attention resources coming from the lowest 20 attention test scores.

Correlation between task performance and workload

For individuals with high attention test scores, there was no correlation between workload and task performance. For individuals with low attention test scores, operators with poor performance tended to indicate having greater overall perceived workload. Their overall perceived workload was correlated with the average reaction time to the red light ($r = .4980$, $p < .0255$), and with the reaction time to the green light ($r = .5329$, $p < .0155$).

Correlation between situation awareness and workload

For both extreme groups, none of the SAGAT elements were significantly correlated with the overall perceived workload. Operator knowledge about system aspects had no apparent association with overall perceived workload.

Correlation between rated situation awareness and workload

For both extreme groups, there was no correlation between workload and rated SA.

Results for individuals with extreme memory scores

A multiple regression analysis was conducted to identify which memory test scores affect overall perceived workload. The regression showed that: (1) a memory span of Numbers, (2) an average percentage of recalling the 7th item from a list of ten, and (3) an average percentage of correctly identifying line 3 were the three variables with the highest effect on perceived workload. Two extreme groups of individuals were then obtained in the same manner as for the attention limitations.

Correlation between task performance and workload

For individuals with low memory test scores, performance was not significantly correlated with the overall perceived workload. For individuals with high memory test scores, proportion of scale misses was associate with low perceived workload ($r = .4821$, p , .0313). Operators with higher proportion of scale misses tended to report higher overall perceived workload.

Correlation between situation awareness and workload

For individuals with high memory test scores, the overall perceived workload was correlated with the knowledge regarding the lights ($r = -.451$, $p < .025$). Operators with more accurate knowledge regarding the lights tended to report lower overall perceived workload. For individuals with low memory test scores, there was no significant relationship between SA and workload found.

Correlation between rated situation awareness and workload

For individuals with high memory test scores, the overall perceived workload was correlated with self rated SA ($r = -.4640$, $p < .0393$). Operators with higher rated SA tended to report lower overall perceived workload. For individuals with low memory test scores, there was no significant relationship between rated SA and workload found.

Table 1. The significant Pearson correlation coefficients between overall perceived workload (TLX) and: (1) task performance, (2) SA, and (3) Rated SA, among different groups of individuals ($p < .05$)

TLX of Groups of Individuals	**Task Performance**					**SA**		**Rated SA**
	R	G	MS	D_A	D_B	L1	P_sup	
Entire participant pool								-.3445* .004**
Individuals with low Attention test scores	.4980* .0255**	.5329* .0155**						
Individuals with high Attention test scores								
Individuals with low memory test scores								
Individuals with high Memory test scores			.4821* .0313**			-0.4993* .025**		-.4640* .0393**

** Correlation*
*** Significant Probability*
R : Average reaction time to the red light
MS : Proportion of scale misses
MG : Proportion of green light misses
D_A : Average deviation in tank A
D_B : Average deviation in tank B
L1 : Knowledge of the current status of the lights
P_sup : Knowledge regarding the pumps

Discussion

Entire participant pool

Operator perceived workload was not found to significantly correlated with task performance. This result was different from the findings found in Raby and Wickens (1994) and Hancock (1996). Two reasons were hypothesized to explain the dissociation between task performance and overall perceived workload.

First, subjects were in different motivation states (Salvendy, 1997 page 434). The participating subjects received a free credit for one homework assignment, which would yield about one percent of their total course grade. Although it was not explicitly stated, the subjects were aware that failing to provide maximum effort in this experiment would not result in any negative consequences. The limited compensation and lack of a penalty for poor performance may discourage the participants from giving the best effort. Having little motivation to meet their goal may alter their performance, as well as reduce their perceived workload as the quality of their performance was not important.

Secondly, no feedback was provided on performance. Becker, et al (1991) found that feedback reflecting good performance could reduce perceived workload, while feedback reflecting bad performance could increase perceived workload. In this study, it may have been possible that while performing the monitoring task, the participant did not realize what he/she was missing, and if he/she needed to respond to the changes faster. Furthermore, feedback from the resource management task performance was not available. Although subjects were informed that the farther the deviations of the fuel from the desired level, the more dangerous the system became, the subjects did not know the exact severity levels.

Having no feedback on performance therefore could keep the perceived workload constantly low throughout different experimental conditions.

Surprisingly, overall perceived workload could not be used to indicate how well the operator knew about system aspects. None of the SAGAT elements were significantly correlated with operator perceived workload. Although this result supported several previous findings that there is no correlation between workload and situation awareness, the existence of such a correlation is still disputable.

Hallbert (1997) has pointed out an interesting point regarding the correlation between workload and situation awareness. He found that a sudden change of workload could cause a decrease of SA, however operators could later regain their SA while the workload remains high. In our study, the operator situation awareness was measured at randomly selected times. Some time points were close to system element failures while others were not. The failures corresponded to the sudden change of workload. Measuring situation awareness at different times from the point of sudden workload change could alter the results. At those instants in time, the operator may (or sometimes may not) have regained his/her situation awareness.

Another reason is that, as mentioned earlier, there was no pressure applied on the participants to achieve their best. This condition could affect the effort spent on the task, and could possibly affect situation awareness. As expected, the operator with low self rated situation awareness had low perceived workload.

Individuals with similar attention and memory limitations

The results showed that individual differences in terms of the limitations of attention and memory affected how operators rated their perceived workload. This hypothesis is supported by the stronger correlation coefficients obtained from the groups of individuals with similar memory and attention limitations than compared to the entire group of the participants. In addition to the stronger correlation coefficients, using attention and memory limitations to divide individuals into four extreme groups revealed several interesting links among the workload, task performance, and situation awareness.

First, there was no apparent association between perceived workload and task performance among individuals with high attention test scores, while some performance components were found to be correlated with perceived workload among individuals with low attention test scores. These results are supported by Endsley (1996 in O'Brien and Charlton, 1996) and Endsley (2000 in Endsley and Garland, 2000) that perceived workload fluctuates and affects task performance especially when it exceeds human capacity. Attention demands for the MAT task may have exceeded the attention capacity of those with low attention scores, resulting in more obvious relationships.

Second, unlike other groups, individuals with higher memory test scores showed an association between perceived workload and situation awareness. Individuals in this group reported having low workload while they were achieving high situation awareness regarding the light status. This relationship could be caused by the effort needed to maintain situation awareness at a certain level. Since there were only two lights, very little effort was needed to reach a certain level of situation awareness. Therefore, operators reported having low workload, while having high SA (Moray, Dessouky, Kijowski, and Adapathya, 1991).

Third, there was no relationship found between self rated SA and perceived workload among individuals with similar attention limitations, neither groups with high or low attention test scores. It could be hypothesized that individuals with low attention test scores perceived the increase of workload while trying to maintain the situation awareness constant at a certain level (Endsley,1993; and Borrensen, Bateman, and Malzahn, 1988). On the contrary, individuals with higher attention test scores could have increased their level of situation awareness while the perceived workload remained constant.

Finally, there was no relationship found between self rated SA and perceived workload among individuals with low memory scores. A similar hypothesis could be stated that individuals with low memory test scores could increased their level of situation awareness while the perceived workload remain constant.

In summary, this study found significant impacts of attention and memory limitations on correlation between perceived workload, task performance, and situation awareness. They resulted in not only stronger correlations amongst workload, task performance, and situation awareness, but also different set of correlated components. Therefore, subjects who participate in system evaluations should closely represent the user population, especially in terms of attention and memory limitations. Future research may extend the current study by including other individual factors such as ages, genders, education level, or related experience.

References

Becker, A.B., Warm, J.S., and Dember, W.N. (1991). Effects of Feedback on Perceived Workload in Vigilance Performing. In *Proceedings of the Human Factors Society 35th Annual Meeting*, (pp. 1491-1494). Santa Monica, CA: Human Factors and Ergonomics Society.

Borrensen, K.J., Bateman, R., and Malzahn, D. (1988). Techniques of Subjective Assessment: A comparison of the SWAT and Modified Cooper-Harper Scales. In *Proceedings of the Human Factors Society 32nd Annual Meeting*, (pp. 155-159). Santa Monica, CA: Human Factors and Ergonomics Society.

Comstock, J.R., Jr., and Arnegard, R.J. (1992). *The Multi-variable Attribute Task battery for Human Operation workload and Strategic Behavior Research.* NASA Technical Memorandum 104174. Hampton VA: National Aeronautics and Space Administration, Langley Research Center.

Endsley, M.R., & Garland, D.J. (2000). *Situation Awareness Analysis and Measurement,* Lawrence Erlbaum Associates. New Jersey.
Endsley, M.R. (1993). Situation Awareness and Workload: Flip side of the Same Coin. In *Proceedings of the 7th International Symposium on Aviation Psychology*, (pp. 906-911)

Endsley, M.R. (2000). *Situation Awareness, Automation & Free Flight* http://vega3.uneec.eurocontrol.fr/endsley.htm, Date accessed: 11/29/2000.

Endsley, M.R., Selcon, S.J., Hardiman, T.D., and Croft, D.G. (1998). Comparative Analysis of SAGAT and SART for Evaluations of Situation Awareness. In *Proceedings of the Human Factors Society 42nd Annual Meeting,* (pp. 82-86). Santa Monica, CA: Human Factors and Ergonomics Society.

Hallbert, B.P. (1997). Situation Awareness and Operator Performance: Results from Simulator-Based Studies, *IEEE,* 18-1 - 18-6.

Hancock, P.A. (1996). Effects of Control Order, Augmented Feedback, Input Device, and Practice on Tracking Performance and Perceived workload. *Ergonomics*, 39(9), 1146-1162.

Heers, S.T., and Casper, P.A. (1998). Subjective Measurement Assessment in a Full Mission Scout-Attack Helicopter Simulation. In *Proceedings of the Human Factors Society 42nd Annual Meeting,* (pp. 26-30). Santa Monica, CA: Human Factors and Ergonomics Society.

Moray, N., Dessouky, M.I., Kijowski, B.A., and Adapathya, R. (1991). Strategic behavior, workload, and performance in task-scheduling. Human Factors, 33, 607-629.

Nygren, T.E., Schnipke, S., and Reid, G. (1998). Individual Differences in Perceived Importance of AWAT Workload Dimensions: Effects of Judgement and Performance in a Virtual High Workload Environment. In *Proceedings of the Human Factors Society 42nd Annual Meeting,* (pp. 816-820). Santa Monica, CA: Human Factors and Ergonomics Society.

O'Brien, T.G., and Charlton, S.G. (1996). *Handbook of Human Factors Testing and Evaluation*, Lawrence Erlbaum Associates. New Jersey.

Raby, M., and Wickens, C.D. (1994). Strategic Workload Management and Decision Biases in Aviation. *The International Journal of Aviation Psychology*, 4(*3*), 211-240.

Salvendy, G. (1997). *Handbook of Human Factors and Ergonomics* 2nd Edition. John Wiley&Sons. New York.

Selcon, S.J., Taylor, R.M., & Koritsas, E. (1991). Workload or Situation Awareness?: TLX vs SART for Aerospace Systems Design Evaluation, *Proceedings of the Human Factors Society 35th Annual Meeting*, 62-66.

Chapter 8. General Conclusions and Discussion

Conclusions and Discussion

Previous psychological studies have demonstrated the existence of the natural limitations of human attention and memory. The attention limitations included: (1) time requirement to reorient attention, reorientation may be expedited by providing spatial cue (2) the limited attention resource available to process targets presented in rapid series, (3) time requirement to search for feature targets and conjunctive targets, (4) automatic behaviors slow down information processing, and (5) location irrelevancy between the stimuli and the response slow down the response identification process. Memory limitations include: (1) memory capacity, (2) false memory, (3) ability to recall items presented in a list is limited, and, (4) the ability to identify uni-dimensional items is limited.

The first paper (Chapter 3) collected information regarding natural attention and memory limitations of undergraduate engineering students. The attention and memory limitations of engineering students participated in this study were similar to that of non-engineer participants of previous studies, yet engineering students tended to better overcome most of such limitations. For example, engineering students could quickly reorient the attention to locate the target regardless whether a spatial cue was provided or not. The effect of automated behavior, and the location irrelevancy between the stimuli and the response were found had less impact on engineering students. Engineering students were also found to have larger memory span, especially in remembering digits. Engineering students were poorer in identifying items at the beginning and at the end of the list, yet they were better identifying items in the middle of the list. Engineering students in this study however, were less effective in overcoming the effect of false memory. The collected data suggested that education background might indicate how well individuals could overcome such limitations in comparison to less educated individuals.

The second paper (Chapter 4) examined the effect of system factors such as automation, pace, and task duration on situation awareness and task performance of operators of a simulated complex system. Automation level and pace of work showed a stronger affect on task performance than task duration did. Automation was the strongest factor affecting situation awareness, while pace yielded lesser effects, and no effect from extended task duration.

The third (Chapter 5) and fourth (Chapter 6) papers examined the effect of individual differences on situation awareness and task performance of operators of a simulated complex system. These two papers showed that individual differences significantly correlated with situation awareness and task performance. The limitations of human attention had a significant correlation with situation awareness, and task performance. Individuals who could score higher on the Spatial Cuing, Stroop Effect, and Simon Effect tests tended to have better task performance and higher situation awareness when performing a simulated complex task under various system conditions. However, higher scores on the Attentional Blink yielded negative relationship, and Visual Search tests yielded mixed relationships.

The effects of memory limitations on situation awareness and task performance of operators during a simulated complex task were similar to the effects of attention limitations. Individuals who could score higher on each of the four memory limitation tests tended to have better task performance when performing a simulated complex task under varying system conditions. Individuals with higher scores on the False Memory, Serial Position, and Absolute Identification tests tended to have higher situation awareness.

Individual differences not only affected the situation awareness and task performance of the operators of a simulated complex system, but also on how the operators felt toward the system they were interacting with. The forth paper (Chapter 7) found significant impacts of attention and memory limitations on the correlation between perceived workload and several aspects situation awareness, as well as a correlation between perceived workload and several aspects of task performance of operators of a simulated complex system.

Similarity in attention and memory limitations resulted in not only stronger correlation coefficients amongst workload, task performance, and situation awareness, but also different set of correlated components.

Clear implications can be drawn from the results obtained from the current study. It is concluded that both system factors and individual factors must be taken into account in the early stage of system design, as well as system evaluation.

In the design stage, the level of automation, pace of work, and task duration must be carefully incorporated. The information regarding attention and memory limitations should also be taken into account to minimize their effect on situation awareness and task performance. The designer should ensure that the attention and memory required to complete the task will not exceed the memory and attention limitations of the user population.

During the system evaluation stage, the prototype should precisely represent the real system. Automation level, pace of work, and task duration must be maintained at the operational level expected to be used in the real system. Speeding up a test session by raising the pace of work, and or decreasing the test duration may lead to invalid conclusions. In addition, subjects who participate in system evaluations should closely represent the user population, especially in terms of attention and memory limitations.

Recruiting operators who could better overcome the limitations of attention and memory may be beneficial in achieving high situation awareness and better task performance. Proper systematic training may also help the operator to overcome both task factors and individual factors so that could high situation awareness and better task performance can be achieved.

Recommendations For Future Research

The current study has indicated several important effects of system factors and individual differences on situation awareness, task performance, and perceived workload. However, more investigation is required to confirm the current findings, as well as to open opportunity for further discovery.

First, extended the survey of attention and memory limitations to students from various educational departments (i.e. business school, engineering school, and social science school). The data from the various schools will then be compared to verify the differences of such limitations for individuals having different educational backgrounds.

Second, to determine the correlation between individual differences and situation awareness, individual differences and task performance, and the effect of individual differences on perceived workload by incorporating different interface designs. Previous studies have shown that interface design strongly affects situation awareness and task performance. Therefore, adding an interface design factor would ensure the effect of individual differences on situation awareness and on task performance.

Third, incorporating various forms of information into the simulated task. Due to the advent of increased availability of technology, several forms of information such as visual and auditory now can be simultaneously presented to the operators of complex systems. The current study included only visual information, however future studies should incorporate new form of information, such as audio.

Fourth, incorporating different forms of task settings, such as automation failure warning light, level of severity incident if the operator fails to meet the goals, or compensation levels into the simulated task.

Fifth, future research could include a motor skill battery test in the experiment. Situation awareness is considered to be a stage before decision making and performing actions, therefore motor skill is considered separately. However, forming situation awareness, making a decision, and performing an action are that stages that highly related. Motor skills therefore could indirectly affect situation awareness, and task performance. Therefore, incorporating a motor skills assessment could make the conclusion more applicable.

Sixth, examine the effect of individual differences on situation awareness and task performance of operators of a high fidelity simulated task. Results would verify the effect of individual differences on situation awareness and task performance.

In summary, the current study could be extended in various ways such as investigating the attention and memory limitations of subjects from various educational or professional backgrounds; and incorporating different interface designs, different forms of information, a motor skill test, different task settings, or a high fidelity simulated task.