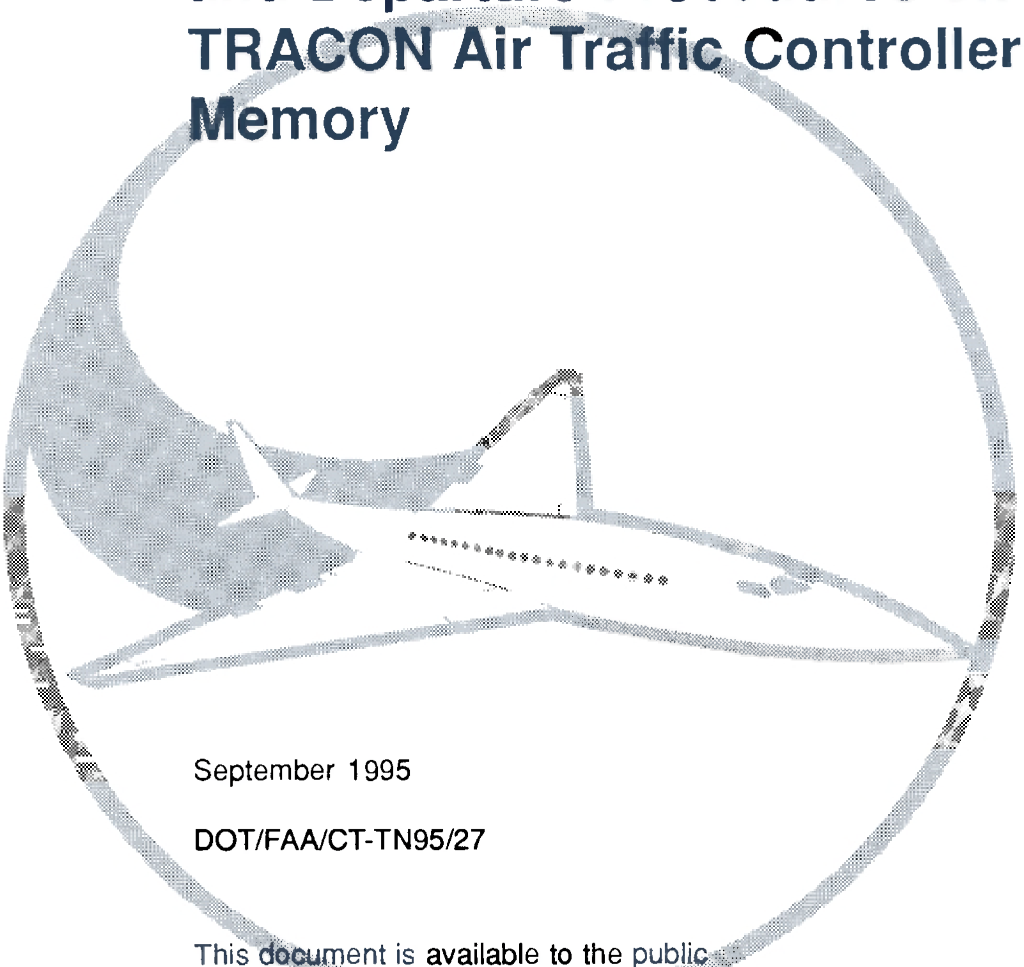
ote technical note techn

The Effects of Structured Arrival and Departure Procedures on TRACON Air Traffic Controller Memory



September 1995

DOT/FAA/CT-TN95/27

U.S. Department of Transportation
Federal Aviation Administration

Technical Center
Atlantic City Airport, NJ 08405

NOTICE

Technical Report Documentation Page

1. Report No. DOT/FAA/CT-TN95/27	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle The Effects of Structured Arrival and Departure Procedures on TRACON Air Traffic Controller Memory and Situational Awareness		5. Report Date September 1995 6. Performing Organization Code ACT-530
7. Author(s) Randy L. Sollenberger, Ph.D. (PERI) and Earl S. Stein, Ph.D ACT-530		8. Performing Organization Report No. DOT/FAA/CT-TN95/27
9. Performing Organization Name and Address Federal Aviation Administration Technical Center Atlantic City International Airport, NJ 08405		10. Work Unit No. (TRAIS) 11. Contract or Grant No. F2202J
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Aviation Administration Office of Aviation Research Human Factors Division Washington, DC 20591		13. Type of Report and Period Covered Technical Note 14. Sponsoring Agency Code AAR-100
15. Supplementary Notes		

16. Abstract

Air traffic control (ATC) is conducted by men and women of the Federal Aviation Administration's (FAA) air traffic service. Controllers do an excellent job of keeping aircraft separated and safe. However, they do make mistakes. Many of these errors are caused by the limitations of working memory, which controllers continuously use to maintain situational awareness (SA).

An experiment was conducted at the FAA Technical Center Human Factors Laboratory to examine the potential benefits of a memory aiding concept on controller performance, SA, and workload. The advanced use of Standard Terminal Arrival Routes (STARs) and Standard Instrument Departures (SIDs) were selected as the memory aids for testing. These specially-designed STARs and SIDs were intended to simplify the controller's task and allow more time for planning and monitoring aircraft. A new high-fidelity ATC simulator was used which allowed controllers to work under extremely realistic conditions. Sixteen controllers from Atlantic City TRACON participated and worked scenarios consisting of low and high traffic volumes both with and without the memory aids. Controllers' actions and aircraft data were recorded during each scenario and used to evaluate ATC performance. Other evaluation methods included the Air Traffic Workload Input Technique and a modification of the Situational Awareness Global Assessment Technique.

The results indicated that the memory aids decreased both the number of ground-to-air transmissions and handoff errors. Controller workload and SA were primarily determined by the traffic volume and were not affected by the memory aids. A final debriefing with controllers suggested several ways the memory aids and SA technique could be improved.

17. Key Words Air Traffic Controller Memory Situation Awareness Workload Memory Aids Performance		18. Distribution Statement This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 104	22. Price

Form DOT F 1700.7 (8_72) Reproduction of completed page authorized

ACKNOWLEDGEMENTS

We would like to thank Stan Gromelski (PERI) for his invaluable air traffic control advice and for serving as our subject-matter expert and observer in this research. We wish to acknowledge the contributions of Ed Buckley (ACT-530), Paul Stringer and Jerry Guttman (PERI), and Jamie Fischer (CTA) for editing drafts of this research paper. We are grateful to Richard Algeo (TRI) for the development of the data reduction and analysis module used in this research and Kathy Mann (PERI) for producing summary reports of the data. We would also like to thank Jerry Guttman, and Tom Zurinskas and Joe Galushka (ACT-530), Albert Macias (ACT-510), Jamie Fischer, and Matthew Platshorn (ACT-530 co-op student) for their contributions as pseudo pilots and data scoring assistants. This research would not have been possible without the technical support of Vince Lasewicz (ACT-540), Dennis Filler, Mike Pomykacz, and Mary Delemarre (ACT-510), George Rowand (SRC), and Nick Marzelli (CTA).

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

EXECUTIVE SUMMARY

Controllers are required to process vast amounts of information in order to conduct air traffic safely and expeditiously. Since most air traffic information is constantly changing, working memory and situational awareness (SA) are critically important for controllers. Human working memory has a limited capacity and is often thought to be a contributing factor to operational errors in air traffic control (ATC). One approach to reducing the incidence of errors is to enhance working memory by providing memory aids to controllers. The primary purpose of this research was to develop and evaluate memory aids intended to improve air traffic controller effectiveness by:

a. Improving working memory and reducing memory-related errors.

b. Investigating the effects of the proposed memory aids on controller SA and workload.

The memory aids investigated in the experiment were specially-designed arrival and departure procedures based on Standard Terminal Arrival Routes (STARs) and Standard Instrument Departures (SIDs). However, the procedures used in the present study differed from basic STARs and SIDs in several important ways. Unlike basic STARs and SIDs, the experimental arrival and departure procedures consisted of a sequence of fixes and altitude changes at designated points to direct arrival and departure aircraft entirely through the terminal environment. These procedures were intended to reduce the need for controller communications and structure (or standardize) the flow of traffic arriving from and departing to many different locations. Since fewer communications were required, controllers could theoretically devote more time to scanning the radar display, reviewing flight progress strips, and performing other activities that should increase their SA. Also, once the experimental arrival and departure procedures have become familiar to controllers, they can serve as a "schema" or "mental model" for organizing and remembering aircraft information.

Sixteen air traffic controllers from the Atlantic City Terminal Radar Approach Control (TRACON) participated in the study. The experiment was conducted at the Federal Aviation Administration (FAA) Technical Center's Human Factors Laboratory at the Atlantic City International Airport in New Jersey. The experimental apparatus consisted of a high-fidelity ATC simulator with voice communication equipment to allow controllers to issue commands to remote pseudo pilots. Each controller performed 8 different scenarios over 2 days of testing. On one of the days, controllers used their own techniques without any special instructions from the experimenters. On the other day, participants used the specially-designed arrival and departure procedures as memory aids while controlling traffic. Before working the actual scenarios, participants received a 1-hour training session to become familiar with the experimental arrival and departure procedures. Controllers performed two low traffic scenarios and two high traffic scenarios on both days of testing to evaluate the memory aids under different traffic conditions. Low traffic scenarios consisted of 14 aircraft appearing within the 30-minute duration of each scenario, and high traffic scenarios consisted of 23 aircraft appearing within the same 30-minute time period.

Data regarding ATC performance, SA, and workload were collected during the simulation. The performance measures included the number of conflict errors, handoff errors, controller assignments, controller transmissions, aircraft density, and flights completed. Controller workload was assessed using the Air Traffic Workload Input Technique (ATWIT). The ATWIT technique consisted of collecting participants' ratings of workload as they controlled traffic. SA was evaluated using a modification of the Situational Awareness Global Assessment Technique (SAGAT). The SAGAT technique consisted of having participants answer questions about the current situation without viewing displays. In this study, the method was used in two different phases during each scenario to collect different information. In the information recall phase, controllers were asked questions about aircraft such as current altitude, heading, and airspeed, and most recently assigned altitude, heading, and airspeed, as well as flight plan information. In the spatial location phase, controllers were asked to locate aircraft on a map of the radar display. In both phases, scoring was accomplished by awarding 0, 1, 2 or 3 points, depending upon the accuracy of controllers' responses.

In addition to the previously described measures, several questionnaires were used to collect subjective ratings from participants and an expert observer. First, an Initial Questionnaire requested background information from each participant. Next, an Information Importance Form asked participants to make importance ratings for different types of ATC information. As controllers worked each scenario, an ATC expert made performance and workload ratings on an Observer Checklist. After each scenario was finished, controllers made performance, SA, and workload ratings in a Post-Scenario Questionnaire. After all the scenarios were finished, the participants were asked to complete the Information Importance Form again to determine the consistency of their previous importance ratings. A Final Questionnaire required controllers to make ratings on the effectiveness, usability, and acceptability of the memory aids.

As expected, the results indicated that the memory aids greatly decreased the number of ground-to-air transmissions, including both altitude and heading assignments. Also, the memory aids reduced the number of handoff misses, but did not affect the number of conflict errors. Controllers' and the expert observer's ratings of performance indicated that the memory aids appeared to slightly increase performance in low traffic scenarios, but seemed to slightly decrease performance in high traffic scenarios. Controllers' ratings of SA were higher with the memory aids, but the SAGAT measures indicated that the memory aids had no effect on SA. Contrary to expectations, controllers' ratings of workload were not affected by the memory aids. Overall, the memory aids did not improve controller performance, increase SA, or reduce workload as much as expected. Controllers' comments indicated that the memory aids may have been more effective if they had more time to learn the experimental arrival and departure procedures. Although some disadvantages were mentioned, most controllers felt the memory aids were helpful, and the ratings of effectiveness, usability, and acceptability were favorable. Controllers' comments suggested several ways that the present SAGAT procedure could be improved in future research.

1. INTRODUCTION.

1.1 PROBLEM STATEMENT.

Memory is an elusive construct that can have a major impact on human performance in complex systems. To the extent that an operator depends on his/her memory, there is always a potential for either forgetting necessary information or ineffectively coding and storing it. In both the near and far terms, it really does not matter why the information was not there when it was needed. What matters is that the lack of accurate information led to an ineffective decision or a maladaptive behavior that had a negative impact on system performance. Whatever human memory is, it both extends and limits human performance.

Air traffic control (ATC) is a system that is not very error tolerant. Small mistakes can lead to costly situations, which fortunately, in most cases, are recoverable. It is not unusual for a controller who has made a mistake to admit that he/she forgot something. What they have forgotten, or have not successfully retrieved, is a piece of data about the way things are (the current situation) or the way things will be if they had done what they intended to when the plan for an aircraft or situation was established and stored in memory.

Warm and Dember (1986) discussed levels of alertness and what we would call, today, situational awareness (SA). They expressed concern that systems are not universally well designed to foster the level of attention and comprehension necessary for continuous successful operations. Their views apply to the complex world of ATC as well as to other high-reliability organizations. The research reported here was begun in an attempt to improve controller performance, given the ATC system as it exists today.

1.2 ASSUMPTIONS AND GOALS.

Controllers work in a very dynamic world in which things are constantly changing, and these changes impose themselves on what the controller is currently trying to accomplish. Keeping up with cognitive processing or SA is critical, but not always successful. This multi-year program of research in controller memory enhancement began in 1989 with a few assumptions and some very specific goals. Assumptions included the belief that the program would not solve all controllers' problems and that there were other dynamics that lead to errors besides memory lapses. It was also assumed that controllers needed help in the here-and-now rather than in the sometime future, when hardware and software may be very different. This meant that the memory program focused on the current state-of-the-art and on what could be done to reduce the probabilities of memory lapses in the system as it exists.

The specific goal of the research reported in this document was to examine the feasibility of some very basic memory aids in terms of their effect on controller performance and SA. Measures of SA were to be used as indicators of the controllers' current working memory store. These concepts will be more clearly defined in latter sections of the report.

ATC is a human-centered system supported by rapidly aging technology. Controllers are exceptional human beings who make the system work despite its and their limitations. It is the safest system of its level of complexity in the world. "Human factors issues have become the last frontier of aviation safety. If significant improvements are to be made in the technology of aviation safety, they must come in the area of human factors" (American Psychological Association, Human Factors and Ergonomics Society, and Federation of Behavioral, Psychological, & Cognitive Sciences, 1994).

Technological advancements continue to be made to meet the ever increasing demands for service. Numerous hardware and software tools have been developed to improve the safety and productivity of the ATC system. Yet, despite all the automated and semi-automated aids available, controllers must still make use of the same skills they relied on previously. That is, they must still plan, organize, scan, decide, and remember. As they do these things, often simultaneously, there remains the possibility for human error. An FAA task force (Operational Error Analysis Work Group, 1987) studied the frequency and possible reasons for operational errors in ATC. An operational error represents a mistake made by a controller, which fortunately leads, in most cases, only to a minor violation of airspace separation standards. The task force identified memory lapses as one major source of such errors.

While controllers add flexibility and adaptability to the system, they also add the potential for error, as Senders and Moray (1991) have described, "All of us have experienced human error. When we interact with machines or complex systems, we frequently do things that are contrary to our intentions. Depending on the complexity of the system and the intentions of the people interacting with it, this can be anything from inconvenience (often it is not even noticed) to a genuine catastrophe."

Memory requirements in ATC are continuous. Air traffic controllers are surrounded by sources of information from which they must select the most critical components. They must then code and store this data. However, this is not always done effectively. One of the most common expressions uttered by controllers who have made an operational error is, "I forgot!" When information is not effectively stored in a timely fashion, it is either not available when needed or it will be retrieved incorrectly. Without ready access to information, SA may not be adequately maintained, and a crisis may result.

In 1993, there were 764 controller operational errors in the United States (Federal Aviation Administration, 1994). This represented a slight increase from 738 errors the previous year. The FAA is constantly working towards eliminating any such errors. Efforts to enhance controller memory are designed to assist the FAA in reaching this goal.

1.3 MEMORY RESEARCH LITERATURE

In an early comprehensive study of controller errors, Kinney, Spahn, and Amato (1977) analyzed FAA reports and developed eight categories of errors. These included: controlling in another's airspace, timing and completeness of flight data handling, inter-positional coordination of data, use of altitude on the display, procedures for scanning and observing flight data, phraseology and use of voice communications, use of human memory to include relying on recall in a noisy environment, and dependence on automatic capabilities.

The FAA uses a different set of categories to classify operational errors, and they include: radar display, communication, coordination, aircraft observation, data posting, and position relief

(Federal Aviation Administration, 1988). By far, the most frequent source of errors identified by the FAA was in a subclass of "radar display: the misuse of data." This category implies that information was available and was either misinterpreted or inaccurately stored in working memory. It is clear that most errors are not induced by taskload and, in fact, most often occur during low to moderate levels of environmental demand. This finding transcends national boundaries and has been demonstrated in analyses of errors conducted, for example, by Transport Canada (Stager & Hameluck, 1990). These authors suggest that previous taxonomies of errors have been incomplete and may have missed information processing failures that subsequently led to inappropriate actions.

The rate at which information flows through the ATC workstation cannot be completely controlled (Sperandio, 1971; Kirchner & Laurig, 1971; Thomas, 1985). The amount of information and the speed with which it can be processed are limited (Finklman & Kirchner, 1980; Spettel & Liebert, 1986; Warm & Dember, 1986). Controllers must, therefore, be able to manage memory successfully in order to select and retain all of the critical elements that confront them.

Opinions concerning the extent of human information processing limits have varied considerably. Miller's (1956) concept that we process about 7 (plus or minus 2) chunks of information at any one time has become accepted doctrine, despite the fact that evidence has shown otherwise under certain conditions (Klapp, Marshburn, & Lester, 1983). The "7 plus or minus 2" view may be relevant for static memory where there are few external sources of interference, but may be too optimistic for dynamic memory which is the reality of most complex command and control systems (Moray, 1986). It is likely that actual working memory is a multi-operational system which includes static memory, dynamic memory, and attentional components (Baddeley, 1986). Long term memory, which contains practically an unlimited storehouse of information gathered over a lifetime, may be an asset or a liability depending on how it assists the management of working memory and SA. This is where the action is in ATC.

Working memory for the air traffic controller is dynamic. In order to manage aircraft, information must be captured and retained for tactical use (3 to 5 minutes) and, secondarily, for strategic planning. Each aircraft's call sign, type, route, and so forth, must be retained for as long as it is under an individual's control, and then discarded. While under control, other information (e.g., altitude, speed, and direction) must be continuously updated and readily accessible so that separation of aircraft can be maintained. Controllers' memory requirements are further burdened by additional demands, such as inclement weather or emergency situations, which may require deviations from the usual expected courses of action.

Flight strips serve in today's system as multifaceted tools for maintaining an ongoing record of events associated with each aircraft. Controllers are required to annotate these strips with changes that they make to the flight plans along with other operational considerations (Federal Aviation Administration, 1989). In addition to writing notes, controllers often rearrange strip placement to act as reminders as to what they have done with aircraft and what they will need to do in the future. The important connection between flight strips and controller memory has been noted by Vortac (1991). This report indicated that memory is essential in understanding the relationship between flight strips and controller performance. However, the relationship may not be obvious.

The value of flight strips has recently been addressed by a study conducted at the Civil Aeromedical Institute (Vortac, Edwards, Jones, Manning, & Rotter, 1992). The authors noted that although controllers often view flight progress strips as unimportant, they do use them. Their study focused on controller behavior in a simulated en route environment and found that note-writing on strips was one of the more frequent activities engaged in by controllers. This group found that as controllers became busier in higher-complexity scenarios, they fell behind in updating the strips. Further, controllers increased the number of requests for information from pilots in higher-complexity scenarios, implying that they could not remember or retrieve all the data they needed.

Technology may have an impact on how operators deal with information. In more automated systems, it is likely that paper flight strips will be eliminated and replaced by electronic media. It has been duly noted that the impact of automating the tasks currently undertaken with paper flight strips must be determined, since their value has been so widely emphasized (Garland, Stein, Blanchard, & Wise, 1992). Hopkin (1991), writing about future automated systems, commented that paper strips may well serve beyond their originally intended purpose. Hopkin has suggested that strip management activities assist in the maintenance of SA, help controllers remember performed and to-be-performed actions, and also help controllers plan strategies for directing traffic. The proposed removal of paper flight strips has raised the concern that controllers will be more likely to lose SA since active involvement with them will be eliminated (Hopkin, 1991; Jackson, 1989).

An important aspect of flight strip management is that it allows controllers to organize information, enabling it to be recalled more efficiently. The relationship between organization and memory has been widely reported in the psychological literature. Those who organize information more extensively have been able to recall more items at the time of testing than those who organize less (Tulving, 1962). Bower, Clark, Lesgold, and Winzenz (1969) found that more words were remembered at the time of testing if they were initially presented according to an organized framework, such as by category (e.g., metals, stones), than if they were presented randomly. Benefits are also observed when the organizational scheme is self-imposed.

Means et al. (1988) studied the way that en route controllers organized aircraft. They observed that controllers recalled aircraft in groups, invariably drawing one group at a time when tested. When asked to name the groups, controllers labeled them in accordance with a specific type of traffic issue (i.e., arrivals or crossing traffic at a specific fix). Geographical proximity played less of a role in grouping than did the interaction and potential conflicts between members of a group. Organization of information has been identified as the one factor which has the greatest probability of improving memory performance in ATC (Vortac, 1991).

Vortac, Edwards, Fuller, and Manning (1994) examined the potential impact of limiting the ability of controllers to organize information by sorting flight strips. They compared the performance of controllers working with limited unmovable strips to that of controllers who could work the way they normally do in today's en route ATC environment. Overall performance measures, surprisingly, did not find differences between the two groups. The authors did find a difference in prospective memory in that controllers with restricted strips recalled more delayed pilot requests and granted them sooner. The authors concluded that restricted strips did not interfere with performance and may reduce controller information processing workload. However, this conclusion should be considered premature. It implies acceptance of the null hypothesis that there are no differences between the groups, a conclusion that every new statistics student is warned against because of the high probability of statistical error. The authors' main contribution may have been in finding interesting ways of looking at prospective memory.

In another experiment, Vortac et al. (1994) examined the impact of flight strip automation on controller efficiency. Efficiency was defined simply as the number of controller actions remaining to clear the airspace when the ATC simulation scenario was stopped after a standard time period. Three levels were used for automation. No automation involved standard flight strips. Partial automation involved automatic strip updating, but the controller could still move and highlight strips. In full automation, the controller had no control over the strips. Again, the researchers found no significant differences between the groups. The authors noted a slight advantage in prospective memory in that controllers using full automation tended to grant a few more delayed pilot requests. Memory for anticipated actions or prospective memory is a critical component of ATC. However, very little research has been accomplished that sheds any light on it.

Activities like note-writing and other flight strip management techniques (e.g., rearranging) may be important to memory for other reasons as well. There is a finite possibility that motoric enactment, the physical manipulation of something like flight strips, may be the key to remembering future planned actions. The benefits of physical activity on memory have been found in other domains. Koriat, Ben-Zur, and Nussbaum (1990) found that performing action phrases such as, "tear up a sheet of paper" and "blow up the balloon," enhanced recall of those phrases. Memory for phrases whose actions were only imagined was not as high. Activity may involve a deeper level of processing, making information more memorable and accessible, as Norman (1992) recently indicated. By comparing drawing to taking a picture of a scene, Norman described that, ". . . the act of drawing requires a degree of concentration and study that intensifies the experience."

Memory issues are involved in virtually every aspect of ATC. So far in this discussion, memory has been treated as if it were an independent construct, which it really is not. The focus on memory in applied settings has changed in recent times and the more modern construct has been referred to as SA. It has not always been clear at what point memory ends and SA takes over. It does seem apparent, however, that memory, in one form or another, is an essential component of SA. Flach (1994) sees the SA construct as a challenge to traditional thinking concerning research in complex person-machine systems. He argues for a more holistic approach which takes into consideration context, workload, and performance as critical issues in systems operations. There have been a number of definitions of SA that have been offered during its relatively short history. Endsley (1989) defines SA as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future."

There appears to be no general agreement on what constitutes SA (Tenny, Adams, Pew, Huggins, & Rogers, 1992). Hitchcock (1993) commented that SA remains largely undefined but "not as a result of a lack of trying." Common elements seem to include the operators understanding of the meaning of events and the ability to anticipate the consequences of actions. Tenny et al. (1992) reviewed some of the theories of SA and human cognitive processing. They conclude that "people can consciously think about one thing at a time." Working memory has limitations and must be supported by long term memory from which it must retrieve essential information after appropriate cuing. This all takes additional processing and effort. Garland, Stein, Blanchard, and Wise (1993) noted that working memory and SA are dependent on long term memory for guidance and organization.

Endsley and Bolstad (1993), in an attempt to clarify the meaning of SA, emphasized the importance of perception for both working and long term memory. As the individual operator constantly updates SA, perception plays a pivotal role. Any inaccuracies, mis-sampling, or mis-allocation of attention can lead to incorrect working memory and subsequent loss of SA. These authors obtained some positive correlations between SA scores and basic perceptual test measures, such as perceptual speed. They cautioned that their data was preliminary and based on small samples of pilots.

Sarter and Woods (1991) have also expressed concern about the lack of agreement concerning what constitutes SA. However, the same issue has been raised with many of the constructs that human factors specialists use every day, such as "workload," and in many respects, "performance." The authors noted that whatever SA is, it results from recurrent situation assessments by the operator. They are limited by working memory and attentional capacity, but also involve perception and pattern matching. Working memory alone is not the same thing as SA.

Sarter and Woods (1991) concluded that definitions of SA should include both current information, which is consciously available in working memory, and information which is available for retrieval when relevant to the situation. Their definition of SA involves "the accessibility of a comprehensive and coherent situational representation which is being continuously updated in accordance with the results of recurring situation assessments."

The evolution of the SA construct has led to the development of measurement concepts to quantify the construct. Admittedly, the majority of this work has focused on aircraft cockpits and the awareness of its pilots. However, as will be seen later in this report, the work done for pilot SA has provided resources to build a foundation in ATC.

Endsley (1990) noted that early work involved fighter pilots and performance in simulations related to kills and losses in dogfights. SA was evaluated very subjectively. She refers to her measurement model as the Situation Awareness Global Assessment Technique (SAGAT). SAGAT involves developing a question set based on potential events in a scenario. Questions are randomly selected from the set. The flight scenario is frozen at a predetermined point in time. The pilot is removed and asked to respond to the questions. The correctness of answers is determined by referring back to what was actually happening at the point of scenario freeze. Scoring requires storing that information in analog or digital fashion so that comparisons can be made post hoc.

Theory and research have been developed and conducted on memory for many years. Work that specifically addresses the needs of air traffic controllers has been a relatively recent event. At least some of this work has been conducted at the FAA Technical Center in Atlantic City, New Jersey.

1.4 MEMORY RESEARCH AT THE FAA TECHNICAL CENTER.

The goals of the FAA controller memory enhancement program have been to conduct research to examine the factors affecting controller memory and to identify and develop memory aids that would assist controllers on the job. Vingelis , Schaeffer, Stringer, Gromelski, and Ahmed (1990) examined theoretical concepts of memory and identified a workable cognitive model for controllers that encompassed memory issues. These authors adapted a model originally developed by Rasmussen (1987), which involves levels of functioning from skill-based to knowledge-based domains. Each of these domains has its own unique sources of memory-induced error. Vingelis et al. defined controller short term or working memory in terms of its functional requirements (including attention and rehearsal), its contents, organizational structure, operational capacity, and limitations.

The authors emphasized the importance of organization in memory. However, their model focused on the more traditional and static concepts in memory, such as the magic number 7 plus or minus 2 (Miller, 1956). They did not consider the increased limitations imposed by the dynamic aspects of ATC and the fact that dynamic memory may be considerably more limited than was originally thought (Garland et al., 1993). In fact, in a dynamic environment, the amount of information an operator can realistically handle may be down to three chunks.

Another component of the memory project involved the development of the Controller Memory Handbook (Stein & Bailey, 1989). This document was created based on the memory literature as applied to person-machine systems. This handbook combined text and cartoon graphics in an attempt to transfer some key principles of memory to the controller community. As a follow-up to the Controller Memory Handbook, copies were sent to a selected number of facilities along with a two-part evaluation questionnaire. In the first part, respondents were asked to rate the handbook. In the second, respondents were requested to describe how they handled memory on the job in their facilities. Results indicated that personnel liked the handbook and found it useful, although some controllers felt that it was too basic to meet their needs. An even more significant finding was the willingness controllers expressed for stating both the nature of their problems and the techniques they used to deal with them (Stein, 1991).

Respondents identified the following causative issues: coordination, attention, distraction, fatigue, change, overload, and position relief briefings. There was some overlap between these results, those of Kinney et al. (1977), and those reported by the Office of Aviation Safety (Federal Aviation Administration, 1987, 1988). Controllers had their own ideas of what worked for them to help reduce the probability of memory lapses. Most of these centered on what they would call "good housekeeping." This implies the use of effective organizational skills and consistent adherence to procedures that controllers are theoretically taught to do, but do not always do in practice.

A two-step research procedure was instituted in 1991 (Gromelski, Davidson, & Stein, 1992). The first step of this procedure was a mail survey of facility managers. This was followed up by face-to-face interviews with 170 controllers at facilities across the continental United States. One of the most significant findings of this study was that controllers were aware of the memory issues and of the aids available within the system as it exists today. However, they resist using them for the same sorts of reasons that lead them to be reluctant to ask for help if they are overloaded. This somehow violates the controller culture. Controllers did indicate that flight strip management activities, including flight strip organization (cocking, tilting) and marking (keeping notes), were the most used memory-aiding techniques. They also agreed that good controllers engage in certain desirable behaviors. They preplan actions, prioritize work sequences, organize aircraft information, anticipate future states or problems, and use effective communication.

Zingale, Gromelski, and Stein (1992) examined some of these concerns in a laboratory setting in which systematic control was possible. These were the first experiments in a series leading to the research described in this report. The authors worked with college students who were studying to be pilots and had no background in ATC. They were taught the principles of control and were tested using a personal computer-based simulation, TRACON II. TRACON II was originally designed as a computer game, and part of the purpose of these studies was to evaluate the feasibility of using it as a tool to study control performance and memory issues.

In one experiment, participants were either encouraged or discouraged from developing operating strategies in advance of controlling traffic. Results indicated that this had little impact on performance. In a second experiment, the availability of planning time prior to working simulated traffic (2 minutes vs. 5 minutes) was tested. This also did not make a difference. In a third experiment, participants were tested for recall of critical flight information after being given the opportunity to mark flight progress strips. Those who used this memory-jogging technique tended to recall more information and perform better in their control duties. Further, those who wrote on the strips reported lower perceived workload at the end of the simulation than those who declined to write. It was recognized that the decision to write or not may well have been associated with what each participant brought with him/her to the study, including basic abilities and self confidence. This study did demonstrate that the use of college students as an analog sample for air traffic controllers was not really viable. It became clear that actual controllers would have to be employed in the future.

Since the results of the third experiment suggested that note-writing on flight strips may be important to memory and performance in ATC, an additional experiment was conducted to directly evaluate the effect of this flight strip management activity on SA and control ability. Maintaining SA involves the ability to access critical information about aircraft, such as aircraft location and current status (e.g., altitude, speed). A loss of SA, forgetting about an aircraft, or what actions need to be taken, can result in serious consequences. The experiment investigated the effect of note-writing on the control performance and SA of actual controllers recruited at the Technical Center (Zingale, Gromelski, Ahmed, & Stein, 1993). SA was assessed by testing each participant's knowledge of aircraft position and the last command issued to each. As in the prior study, TRACON II, an ATC simulator for the personal computer was the testing device.

This study introduced several interesting artifacts which led to questions concerning the potential for TRACON II as a test bed. While SA and performance did not vary in relationship to the test conditions which involved manipulating taskload, a relationship between them and prior video game experience was identified. Controllers with low self-assessed video game experience showed improved SA when they were allowed to write on the strips. Those with high video game experience did as well in the no writing as in the writing conditions. The principle artifact here may well have been how controllers enter commands in the real world as compared to how they do it in TRACON II. In ATC, they give their commands verbally over a radio net. In TRACON II, they do it with a keyboard. This raises questions about automation issues and how future systems will use their human resources. Anticipated improvements involve a great deal more key punching and considerably fewer verbal commands.

Additional results indicated that controllers were less able to remember the call signs of aircraft in this experiment in which control commands were typed, than they are on the job in which commands are issued verbally. Controllers recalled an average of 82 percent of the call signs in this experiment. One participant recalled only 20 percent. This result suggests that critical information may be lost under conditions in which keyboard entries are used to communicate with aircraft, and has important implications for systems requiring keyboard rather than verbal communication. The current study would avoid these artifacts, because it employed moderate to high fidelity simulation in which controllers had the opportunity to do the job the way they normally do it.

2. EXPERIMENT.

2.1 PURPOSE.

The purpose of this research was to develop and evaluate memory aids intended to improve air traffic controller effectiveness by increasing SA and reducing memory-related errors. The primary memory aids investigated were specially-designed arrival and departure procedures. These memory aids can be implemented in the current National Airspace System (NAS) and may be equally useful in the future system. A secondary goal was to investigate the affect of the proposed memory aids on controller SA and workload.

2.2 MEMORY AIDS.

The memory aids evaluated in this study were based on Standard Terminal Arrival Routes (STARs) and Standard Instrument Departures (SIDs). STARs and SIDs are preplanned instrument flight rules (IFRs) published for pilot use and intended to provide transition between the en route and terminal environments. The experimental arrival and departure procedures designed for the present study differed from basic STARs and SIDs in several important ways. Although most STARs and SIDs are rather simple procedures and involve a single navigation fix, the experimental arrival and departure procedures consisted of navigation to a sequence of fixes. Also, most STARs and SIDs do not involve altitude clearances. However, the experimental arrival and departure procedures consisted of multiple altitude changes at designated fixes. Most small airports only have a few STARs and SIDs, while many experimental arrival and departure procedures were defined in the present study. The experimental arrival and departure procedures were intended to reduce the need for controller communications and to direct arrival and departure aircraft entirely through the terminal environment. The arrival and departure procedures were designed to structure (standardize) the flow of traffic arriving from and departing to many different locations. It is important to note that the arrival and departure procedures did not guarantee safe aircraft separation, and many aircraft in the simulation were not using the experimental procedures. Therefore, controllers still had to monitor radar display activity and take appropriate actions when potential conflicts were detected.

Theoretically, the experimental arrival and departure procedures could potentially improve SA for air traffic controllers in several ways. First, it was expected that controllers would need to spend less time exchanging communications with pilots. Therefore, more time could be devoted to scanning the radar display, reviewing flight progress strips, and performing other activities that should increase SA. Once the experimental arrival and departure procedures have become familiar to controllers, they could serve as a "schema" or "mental model" for organizing and remembering aircraft information. A related potential benefit involved the predictability of aircraft movements that derive from knowledge of the arrival and departure procedures. Aircraft following the same experimental arrival or departure procedure share a common flight plan which may serve as a basis for "chunking" the information from different aircraft into a single unit that could be more easily remembered.

2.3 AIRSPACE AND TRAFFIC SCENARIOS.

One of the primary concerns in this experiment was to create a realistic simulation of Atlantic City International Airport (ACY) TRACON for controllers. To achieve this objective, an air traffic control specialist (ATCS) visited the tower and studied ACY operations. The ATCS talked with controllers about normal operating procedures, observed radar displays, and listened to communications while controllers conducted actual air traffic. Also, airspace boundary data, letters of agreement, and flight progress strips of actual air traffic were obtained. This information was used to develop a realistic depiction of ACY airspace and construct realistic traffic scenarios. It was believed that the efforts invested in creating a realistic simulation would further motivate participant controllers and increase the credibility of the research results.

Using the information obtained from the ACY tower, the airspace was constructed with a few minor modifications. Three additional points in space were necessary to define the experimental arrival and departure procedures that would be used. Also, the altitude boundary of ACY airspace was increased from 7,000 feet to 17,000 feet. This was done so that aircraft arriving, departing, and overflying at high altitude would be worked by ACY controllers instead of en route controllers. This modification allowed the arrival and departure procedures to do more work for the participant controllers. Finally, only four airports were represented in the simulated airspace, although there are as many as six airports in the vicinity. However, very little air traffic is associated with either of the two minor airports that were not represented.

Scenarios were constructed that accurately simulated ACY air traffic. Many of the aircraft call signs were familiar to controllers and represented common air carriers that operate in ACY airspace. However, since most of ACY air traffic does not consist of air carriers, the scenarios depicted a large majority of general aviation (GA) aircraft. Most of the aircraft were arriving to and departing from ACY. However, a small proportion of the aircraft used the three minor airports in the region. Aircraft arrived to and departed from the simulated airspace in the same general directions as typical ACY traffic. All scenarios started without any initial aircraft on the radar display. Then, aircraft steadily appeared creating a buildup of aircraft and maintaining this level of traffic until the conclusion of the scenario. The scheduled rate of appearance for aircraft was changed to represent either moderately busy traffic conditions (hereafter referred to as low traffic) or extremely busy traffic conditions (hereafter referred to as high traffic). It was decided that scenarios would represent IFR conditions and therefore, only IFR traffic was scheduled.

3. METHOD.

3.1 PARTICIPANTS.

Sixteen air traffic controllers from ACY TRACON volunteered for this study. Volunteers were assured of their anonymity and confidentiality. All participants were full performance level (FPL) controllers with normal or corrected-to-normal vision and had actively controlled traffic for 12 months prior to the study. An Initial Questionnaire was completed by each controller to describe the background characteristics of participants in this study. Controllers ranged in age from 24 to 40 years old (Mean=33.13, SD=5.38) and ranged in experience from 1 to 20 years of active service (Mean=7.72, SD=6.13). Additionally, controllers provided self-ratings of four personal attributes that could affect simulation performance. Ratings were indicated on a scale ranging from 1 (meaning low/poor) to 10 (meaning high/good) on each question. The attributes included controller skill (Mean=7.50, SD=0.97), motivation (Mean=8.63, SD=1.26), health (Mean=8.88, SD=1.09), and video game experience (Mean=4.88, SD=2.94).

3.2 SIMULATION FACILITY.

The experiment was conducted in the Human Factors Laboratory (HFL) at the FAA Technical Center in Atlantic City, New Jersey. The experimental apparatus consisted of a state-of-the-art controller workstation with a large high-resolution graphics display, voice communication equipment, networked computer resources, and ATCoach simulation software (copyright UFA Inc., 1992). The simulation was conducted by a research psychologist and an ATCS observing the participant in one experiment room. A voice communication link to another experiment room allowed the controller to issue commands to personnel serving as pseudo pilots. Two trained pseudo pilots provided realistic voice feedback to controllers and operated aircraft using simple keyboard commands. Additionally, the pseudo pilots served as ghost controllers to simulate interaction with other controllers. As part of the simulation materials, flight progress strips were printed and time-ordered in a strip bay prior to the start of each scenario. During the simulation, audio-visual equipment was used to film participants. Permanent recordings of the controller's radar display, voice communications, and actions were made for future reference.

3.3 EXPERIMENTAL DESIGN.

The main independent variable of the experiment will be referred to as the MEMORY condition. This manipulation required that controllers perform 8 different test scenarios over 2 days of testing. On one of the days, controllers used their own techniques without any special instructions from the experimenters. On the other day, participants used the specially-designed arrival and departure procedures as memory aids while controlling traffic. In order to evaluate the effectiveness of the memory aids, performance on both days was compared. The second independent variable will be referred to as the TRAFFIC condition. This manipulation involved constructing four scenarios with a low volume of air traffic and four scenarios with a high volume of air traffic. Low traffic scenarios consisted of 5 arrivals, 5 departures, and 4 overflights for a total of 14 aircraft appearing within the 30-minute duration of each scenario. High traffic scenarios consisted of 8 arrivals, 8 departures, and 7 overflights for a total of 23 aircraft appearing within the same 30-minute time period. The experimental design can be summarized as a 2 x 2 within-subjects (or repeated measures) design with the factors MEMORY (no memory aids, memory aids) and TRAFFIC (low, high). A summary of the experimental design is presented in table 1.

TABLE 1. A SUMMARY OF THE EXPERIMENTAL DESIGN

	NO MEMORY AIDS	MEMORY AIDS
LOW TRAFFIC	2 scenarios	2 scenarios
HIGH TRAFFIC	2 scenarios	2 scenarios

The main dependent variables of the experiment can be categorized into three different areas of interest. The first area of interest was ATC performance or system effectiveness. The present experiment used a long list of common performance measures that have been examined in previous ATC simulation research (Buckley, DeBaryshe, Hitchner, & Kohn, 1983). Although all measures were examined initially, this study will report some of the more important results from a much smaller subset of variables. The performance measures included were: the number of conflict errors, handoff errors, controller assignments, controller transmissions, aircraft density, and flights completed. The second area of interest was SA which was assessed using a modification of the Situational Awareness Global Assessment Technique (SAGAT). The two SAGAT variables examined were the accuracy of controllers' responses in the information recall and spatial location phases. The third area of interest was controller workload, which was assessed using the Air Traffic Workload Input Technique (ATWIT). In addition to the measures just mentioned, ATC performance, SA, and workload were assessed using subjective ratings provided by participant controllers in a Post-Scenario Questionnaire. Two final variables included ratings of performance and workload provided by the ATCS in an Observer Checklist. A complete list of all the dependent variables and their acronyms is presented in table 2.

TABLE 2. DEPENDENT VARIABLES WITH ABBREVIATIONS AND DESCRIPTIONS

Objective Performance Variables

Abbreviation	Description
NCNF	Number of Conflicts (less than 3 miles and 1,000 feet separation)
NHOMISS	Number of Handoff Misses (aircraft crossing boundary without being handed off)
NALT	Number of Altitude Assignments
NHDG	Number of Heading Assignments
NSPD	Number of Speed Assignments
NPTT	Number of Ground-to-Air Transmissions
CMAV	Cumulative Average of System Activity or Aircraft Density (number of aircraft within 8 miles of another aircraft)
NCOMP	Number of Flights Completed

Situational Awareness Variables

Abbreviation	Description
SGTIR	Percentage of Points Obtained in the Information Recall Phase
SGTSL	Percentage of Points Obtained in the Spatial Location Phase

Controller Workload Variables

Abbreviation	Description
ATWIT	Air Traffic Workload Input Technique Rating

Post-Scenario Subjective Variables

Abbreviation	Description
PSQWRK	Controller's Workload Rating
PSQPFM	Controller's Performance Rating
PSQSAW	Controller's Overall Situational Awareness Rating
PSQCAL	Controller's Current Aircraft Location Awareness Rating
PSQPAL	Controller's Projected Aircraft Location Awareness Rating
PSQPSV	Controller's Potential Safety Violations Awareness Rating
PSQDIF	Controller's Scenario Difficulty Rating

Expert Observer Variables

Abbreviation	Description
OBSWRK	Expert Observer's Workload Rating
OBSPFM	Expert Observer's Performance Rating

3.4 PROCEDURE.

When controllers arrived at the HFL, they were given a welcome tour of the facility and briefed on how the experiment was going to be conducted, what was expected from them, and their rights as volunteers. Next, controllers completed the Initial Questionnaire and an Information Importance Form. The Information Importance Form required participants to make importance ratings for different types of ATC information. The purpose of this form was to determine what information controllers thought was most important to them for consideration in future memory aids research. Then, participants performed 5 scenarios on each of the 2 testing days. As controllers worked each scenario, an ATCS made over-the-shoulder observations and completed an Observer Checklist. After each scenario was finished, controllers made ratings in a Post-Scenario Questionnaire. At the conclusion of the final day of testing, participants were debriefed as to the expectations of the experiment and given an opportunity to ask any last questions. Finally, participants completed the Information Importance Form again and a Final Questionnaire. The Final Questionnaire required controllers to make ratings on the effectiveness, usability, and acceptability of the proposed memory aids. Also, participants were given the opportunity to make any suggestions or final comments regarding the experiment. The contents of the above questionnaires are presented in appendix A.

The presentation order of scenarios and counterbalancing features of the experimental design are illustrated in table 3. For half the participants (denoted Group A), the first day of testing was without the memory aids and the second day of testing included the memory aids. For the other half of the participants (denoted Group B), the reverse was true — the first day included the memory aids, and the second day was without the memory aids. The first scenario on each day was a 20-minute practice scenario, followed by four 30-minute test scenarios. The four test scenarios were worked so that each of the two low traffic scenarios alternated with each of the two high traffic scenarios. For half the participants in Group A and Group B, a low traffic scenario was performed first, and for the other half of the participants, a high traffic scenario was performed first. Also, the presentation position of any scenario was counterbalanced across participants so that one controller worked the scenario first, another controller worked the scenario second, and so on. As shown in table 3, an important feature of the design to emphasize is that each scenario was performed only once by each controller. Although each scenario was worked with memory aids and without memory aids, different controllers performed these two versions of the same scenario.

A training program was developed to assist controllers in learning the experimental arrival and departure procedures used in the simulation. The training session lasted approximately one hour and consisted of textual presentations, graphical visual aids, and a demonstration scenario. As the ATCS described the memory aids, controllers studied textual and graphical depictions of the fixes and altitude changes associated with the arrival and departure procedures. Since almost all the fixes were familiar to controllers, it should not have been difficult to learn this information.

TABLE 3. PRESENTATION ORDER OF SCENARIOS AND COUNTERBALANCING FEATURES OF THE EXPERIMENTAL DESIGN

Group A Participant	No Memory Aids					Memory Aids				
	P1	L1	H1	L2	H2	P2	L3	H3	L4	H4
2	P1	L2	H2	L3	H3	P2	L4	H4	L1	H1
3	P1	L3	H3	L4	H4	P2	L1	H1	L2	H2
4	P1	L4	H4	L1	H1	P2	L2	H2	L3	H3
5	P1	H1	L1	H2	L2	P2	H3	L3	H4	L4
6	P1	H2	L2	H3	L3	P2	H4	L4	H1	L1
7	P1	H3	L3	H4	L4	P2	H1	L1	H2	L2
8	P1	H4	L4	H1	L1	P2	H2	L2	H3	L3
Group B Participant	**Memory Aids**					**No Memory Aids**				
9	P1	L1	H1	L2	H2	P2	L3	H3	L4	H4
10	P1	L2	H2	L3	H3	P2	L4	H4	L1	H1
11	P1	L3	H3	L4	H4	P2	L1	H1	L2	H2
12	P1	L4	H4	L1	H1	P2	L2	H2	L3	H3
13	P1	H1	L1	H2	L2	P2	H3	L3	H4	L4
14	P1	H2	L2	H3	L3	P2	H4	L4	H1	L1
15	P1	H3	L3	H4	L4	P2	H1	L1	H2	L2
16	P1	H4	L4	H1	L1	P2	H2	L2	H3	L3

P1 and P2 are similar simple practice scenarios
L1, L2, L3, and L4 are similar low traffic test scenarios
H1, H2, H3, and H4 are similar high traffic test scenarios

Then, controllers observed a demonstration of aircraft arriving from different locations using the arrival procedures and aircraft departing in different directions using the departure procedures. The ATCS demonstrated the necessary communications and techniques as he worked the demonstration scenario. Controllers were given an opportunity to work a simple practice scenario and ask questions before continuing with the formal test scenarios. A display describing the experimental arrival and departure procedures was placed nearby as a reminder while controllers worked the scenarios. Since the controllers were already experienced in using basic STARs and SIDs, it was thought that this training would be sufficient to familiarize them with the specially-designed arrival and departure procedures that would be used in the simulation.

The method used to assess SA was a modification of SAGAT (Endsley, 1987, 1988). SAGAT is an objective technique for evaluating SA that has been applied to pilots in fighter combat simulations and is equally applicable to controllers in ATC simulations. The technique consists of interrupting or "freezing" the simulation and having participants answer questions about the current situation without viewing displays. In the present study, the SAGAT method was used at two different times during each scenario to collect different information. In the first instance, referred to as the information recall phase, the interruption occurred at either 17.5, 22.5 or 27.5 minutes into the scenario. At that time, controllers were asked questions about no more than three aircraft on the radar display. After a few preliminary questions, the rest were randomized and presented one at a time on a computer screen for each aircraft that was not yet handed off. The typical format for each question consisted of an aircraft call sign and a question about that aircraft. In the second instance, referred to as the spatial location phase, the interruption occurred at the end of the scenario and controllers were asked to locate no more than eight aircraft that were still on the radar display. Controllers were instructed to locate only the aircraft that were not yet handed off. The procedure consisted of placing the number associated with each call sign in the proper location on a map of the radar display. For a list of the information recall questions and an example of the spatial location map, see appendix B.

A special scoring procedure was developed for both the information recall and spatial location SA data. The scoring system is conceptually similar to that of target-shooting where more points are awarded for bullets striking closer to the "bull's eye." For each information recall question requiring a numeric response, three different point-scoring ranges were defined. If the response was within the closest range of accuracy, 3 points were awarded. If the response was within the middle range, 2 points were awarded. If the response was within the outer range, 1 point was awarded, and a response beyond the outer range was not awarded any points. Some questions required the name of a fix as the appropriate answer and were scored as hit (3 points) or miss (0 points). Different ranges of accuracy were defined for each question and a list of the criteria is provided in table 4. For each aircraft that controllers were required to locate, three different point-scoring ranges were defined as well. If the aircraft was placed within 5 nautical miles (nm) of its actual location, 3 points were awarded. If the aircraft was placed within 10 nm of its actual location, 2 points were awarded. If the aircraft was placed within 15 nm of its actual location, 1 point was awarded, and an aircraft placed beyond 15 nm was not awarded any points. Both an information recall and a spatial location percentage score were calculated by dividing the number of points that was actually obtained by the number of points that could have been obtained and multiplying by 100.

The method selected to assess controller workload was ATWIT (Stein, 1985). ATWIT provides an unobtrusive and reliable means for collecting participants' ratings of workload as they control traffic. In the present study, a touchscreen was used to present the workload rating scale and record the controllers' responses. Controllers were instructed to indicate their current workload level by pressing one of the touchscreen buttons labeled from 1 (indicating low workload) to 10 (indicating high workload). The touchscreen was programmed to request the controllers' input every 5 minutes by emitting several beeps and presenting the rating scale. Participant controllers had 20 seconds to respond by pressing one of the 10 buttons. If they were too busy to respond within 20 seconds, then the maximum workload rating of 10 was recorded by default. In almost every instance, controllers were able to respond within the allotted time and avoid the default rating.

TABLE 4. ACCURACY LEVELS DEFINING THE DIFFERENT SCORING RANGES FOR THE SAGAT QUESTIONS

Question-Description	3-Point Range	2-Point Range	1-Point Range
3-Within 5 nm/below	± 1 aircraft	± 2 aircraft	± 3 aircraft
4-Within 5 nm/above	± 1 aircraft	± 2 aircraft	± 3 aircraft
5-Within 10 nm/same	± 1 aircraft	± 2 aircraft	± 3 aircraft
6-Current altitude	± 1000 feet	± 2000 feet	± 3000 feet
7-Current airspeed	± 20 knots	± 40 knots	± 60 knots
8-Current heading/fix*	± 10 degrees	± 20 degrees	± 30 degrees
9-Assigned altitude	± 1000 feet	± 2000 feet	± 3000 feet
10-Assigned airspeed	± 20 knots	± 40 knots	± 60 knots
11-Assigned heading/fix*	± 10 degrees	± 20 degrees	± 30 degrees
12-Arrival airport	Hit or Miss Scoring		
13-Entry fix	Hit or Miss Scoring		
14-Departure airport	Hit or Miss Scoring		
15-Exit fix	Hit or Miss Scoring		
#-Spatial Location	within 5 nm	within 10 nm	within 15 nm

-----
* Hit or Miss Scoring was used if the participant responded with a fix name

4. RESULTS.

The results of this experiment will be reported in three major sections. The first section will discuss preliminary analyses conducted on the dependent variables. The approach will rely on a correlational analysis to determine the extent to which ATC performance, SA, and workload are related. Since different techniques (or variables) were used to assess performance, SA, and workload, the results of this analysis will determine the agreement between different measures of the same construct. For example, SA was examined using SAGAT (an objective method) and controller ratings (a subjective method). Measurements collected by these techniques should relate (or correlate) very well. However, if the SAGAT variables do not relate to controller ratings, then this might reduce the utility of the objective measures of SA. Also, it is of interest to identify how the different constructs of performance, SA, and workload related to each other. For example, do SA and workload relate to ATC performance?

The second section is the primary concern of this study. The approach will rely on analysis of variance (ANOVA) to evaluate the effectiveness of the proposed memory aids and examine any differences between low and high traffic scenarios. Dependent measures of ATC performance, SA, and workload will be analyzed to determine if the memory aids had any beneficial (or detrimental) effects. Also, traffic volume will be investigated as a potential influence on the effectiveness of the memory aids. For example, the memory aids may improve SA in high traffic scenarios where controllers may be struggling to "keep the picture." However, the memory aids may be less beneficial in low traffic scenarios where it is relatively easy for controllers to maintain SA.

The last section will summarize the feedback that controllers provided about the experiment. The results of the Final Questionnaire and Information Importance Form will be presented and discussed. The Final Questionnaire provided another means for evaluating the proposed memory aids since many of the ratings and comments concerned the effectiveness, usability, and acceptability of the memory aids. Also, the credibility of the results will be discussed through controller ratings and comments about the realism of the simulation and measurement techniques. Finally, the results of the Information Importance Form will be reviewed. The Information Importance Form was included to determine what ATC information controllers thought was most important to their work. It was administered on both days of the experiment in order to determine the reliability of controller ratings. A correlational analysis will be used to determine the consistency of controller ratings on the two days and an examination of the variability of ratings will be examined to determine the consistency among controllers. The information obtained will be used when considering future memory enhancement experiments. For example, future experiments may investigate the potential benefits of new display techniques which emphasize the information that controllers believe to be most important to their job.

A summary and conclusions subsection is included at the end of each of these three major sections. The purpose of these subsections is to present briefly the most important results and discuss their implications for the goals of this research.

4.1 RELATING PERFORMANCE, SA, AND WORKLOAD.

The results of a correlational analysis relating ATC performance, SA, and workload is reported in this section. A correlational analysis is a formal statistical technique for calculating the degree to which two variables relate or "covary." The results of the analysis produce a correlation coefficient (or r-value) which ranges from -1.0 to +1.0 and indicates the strength of the relationship between the two variables. A coefficient of 0.0 means that no relationship exists, while -1.0 and +1.0 indicate perfect relationships. A positive coefficient (or direct relationship) means that as the value of one variable increases the other variable increases, while a negative coefficient (or inverse relationship) means that as the value of one variable increases the other variable decreases. A correlation coefficient is considered to be statistically significant if its absolute magnitude exceeds a given critical value. The critical value depends on the number of degrees of freedom in the experimental design and can be obtained from most statistics textbooks. Usually, a p-value (or significance level) is reported which represents the probability that the calculated coefficient could exceed the critical value by chance alone. Finally, although it is often tempting to conclude that one variable has a causal effect on the other variable, this is not the correct interpretation of a correlational analysis. A third factor may be causing the two variables to increase and decrease together.

The correlational analysis was based upon 128 observations (16 participants times 8 scenarios per participant). Although a complete correlational analysis relating all 20 dependent variables was conducted, only the dependent variables which had statistically significant correlation coefficients will be reported. Also, the correlation coefficient of variables that were expected to be correlated will be presented as well. The Pearson method of calculating a correlation coefficient was used and the critical values were 0.50 and 0.63 (with 14 degrees of freedom) for significance levels of $p<0.05$ and $p<0.01$, respectively. The results of the correlational analysis were organized into groups of related variables and the coefficients are presented in the tables that follow.

The results of the correlational analysis conducted on the performance measures are reported in table 5 (see table 2 for acronym expansion). As shown, there was a high degree of relationship between many of the performance variables. The number of altitude assignments, heading assignments, ground-to-air transmissions, flights completed, and the aircraft density were significantly correlated with each other. However, the number of flights completed was only marginally correlated with the number of heading assignments. These results were not surprising since these five variables are generally related to the number of aircraft and taskload demands of the scenario. However, the number of airspeed assignments did not relate to any other performance measure. This finding can be explained in part by the fact that controllers rarely used speed adjustments during the simulation. Also, the number of conflicts and handoff errors did not relate to any of the other performance measures in this group. However, conflict and handoff errors rarely occurred, which limited the range of these variables and the probability of obtaining statistically significant correlations. One possible interpretation of these results is that the variables that did not correlate may have been measuring different aspects of ATC performance.

Another issue, which can be addressed by the correlational analysis, is the utility of the objective performance measures produced in this simulation. The participants' and expert observer's ratings of performance have face validity and can serve as standard measurements of performance. The objective measures that correlate with these subjective measures gain credibility as indicators of performance or system effectiveness. The results showed that the number of conflicts and aircraft density were correlated with the expert's performance ratings, but not to the participants' ratings. None of the other objective performance measures related to either the participants' or the expert's ratings.

Table 6 shows the results of the correlational analysis conducted on the SAGAT variables and ATWIT. One interest of the present research was to develop and evaluate an objective measurement technique for air traffic controller SA. The results of the correlational analysis indicated that the SAGAT data collected in the information recall phase and the spatial location phase did not correlate with any of the subjective measures of SA or with each other. Also, although it was reasonable to expect a relationship between SA and performance, the SAGAT variables did not correlate with any of the objective performance or with the participants' and expert observer's ratings of performance. These results did not confirm the utility of the present implementation of SAGAT as an objective measure of SA. The results of the analysis indicated that neither of the SAGAT variables correlated with any of the workload variables. The utility of ATWIT as a technique for measuring controller workload can be examined through correlations with the participants' and expert observer's post-scenario ratings of workload. As shown, there

TABLE 5. CORRELATION VALUES FOR THE PERFORMANCE VARIABLES

	NCNF	NHOMISS	NALT	NHDG	NSPD	NPTT	CMAV	NCOMP
NCNF	.	.	.	.	.	.	.	.
NHOMISS	.	.	.	.	.	.	.	.
NALT	.	.	.	0.89**	.	0.64**	0.59*	0.50*
NHDG	.	.	0.89**	.	.	0.64**	0.52*	.
NSPD	.	.	.	.	.	.	.	.
NPTT	.	.	0.64**	0.64**	.	.	0.66**	0.71**
CMAV	.	.	0.59*	0.52*	.	0.66**	.	0.80**
NCOMP	.	.	0.50*	.	.	0.71**	0.80**	
SGTIR								
SGTSL								
ATWIT			0.50*				0.74**	0.54*
PSQWRK			0.56*	0.51*		0.58*	0.83**	0.62*
PSQPFM	-0.25		-0.26	-0.26		-0.32	-0.44	-0.30
PSQSAW								
PSQCAL								
PSQPAL								
PSQPSV								
PSQDIF			0.56*	0.50*		0.62*	0.83**	0.68**
OBSWRK			0.63**	0.54*		0.62*	0.87**	0.67**
OBSPFM	-0.52*		-0.38	-0.34		-0.29	-0.56*	-0.36

* indicates a significant effect, $p < 0.05$
** indicates a significant effect, $p < 0.01$

TABLE 6. CORRELATION VALUES FOR THE SAGAT VARIABLES AND ATWIT

	SGTIR	SGTSL	ATWIT
NCNF			
NHOMISS			
NALT			0.50*
NHDG			0.43
NSPD			
NPTT			0.48
CMAV			0.74**
NCOMP			0.54*
SGTIR		0.32	
SGTSL			
ATWIT	-0.17	-0.33	
PSQWRK	-0.20	-0.26	0.84**
PSQPFM	0.09	0.24	-0.53*
PSQSAW	0.12	0.29	-0.54*
PSQCAL	0.19	0.31	
PSQPAL	0.14	0.29	
PSQPSV	0.13	0.22	
PSQDIF			0.80**
OBSWRK	-0.21	-0.27	0.85**
OBSPFM	0.13	0.21	-0.54*

-----
* indicates a significant effect, $p < 0.05$
** indicates a significant effect, $p < 0.01$

was strong correlations between ATWIT and the other subjective ratings. Further evidence for the utility of ATWIT was provided by correlations with variables that should have logically related to controller workload. The results indicated that ATWIT was strongly correlated with aircraft density as well as moderately correlated with the number of altitude assignments, heading assignments, ground-to-air transmissions, flights completed, and the aircraft density. These five performance measures are good indicators of taskload or the demands that are placed on participants to safely control the amount of traffic in scenarios. Controller workload is a response to taskload demands and is therefore expected to correlate with taskload measurements. The correlations between ATWIT and the previously mentioned performance variables replicates the validation research conducted by Stein (1985). Negative correlations were found between ATWIT and the participants' and expert's ratings of performance. These results suggest a possible relationship between workload and performance — as workload increases, performance decreases.

The results of the correlational analysis conducted on the Post-Scenario Questionnaire are reported in table 7. The questionnaire requested several ratings regarding different aspects of SA, such as an awareness for current aircraft locations, projected aircraft locations, and potential

TABLE 7. CORRELATION VALUES FOR THE POST-SCENARIO QUESTIONNAIRE VARIABLES

	PSQWRK	PSQPFM	PSQSAW	PSQCAL	PSQPAL	PSQPSV	PSQDIF
NCNF							
NHOMISS							
NALT	0.56*						0.56*
NHDG	0.51*						0.50*
NSPD							
NPTT	0.58*						0.62*
CMAV	0.83**						0.83**
NCOMP	0.62*						0.68**
SGTIR			0.12	0.09	0.14	0.13	
SGTSL			0.29	0.31	0.29	0.22	
ATWIT	0.84**	-0.53*	-0.54*	-0.47	-0.48	-0.45	0.80**
PSQWRK	.	-0.57*	-0.62*	-0.44	-0.50*	-0.51*	0.93**
PSQPFM	-0.57*		0.79**	0.69**	0.68**	0.63**	-0.59*
PSQSAW	-0.62*	0.79**		0.78**	0.76**	0.68**	-0.61*
PSQCAL	-0.44	0.69**	0.78**		0.74**	0.55*	-0.45
PSQPAL	-0.50*	0.68**	0.76**	0.74**		0.73**	-0.49
PSQPSV	-0.51*	0.63**	0.68**	0.55*	0.73**		-0.47
PSQDIF	0.93**	-0.59*	-0.61*	-0.45	-0.49	-0.47	
OBSWRK	0.85**	-0.57*	-0.59*	-0.47	-0.49	-0.48	0.85**
OBSPFM	-0.57*	0.51*	0.54*	0.40	0.46	0.56*	-0.54*

* indicates a significant effect, $p < 0.05$
** indicates a significant effect, $p < 0.01$

safety violations, as well as an overall rating of SA. The analysis indicated that all four SA ratings were strongly correlated. These results suggest that the different questions were not measuring independent aspects of SA, and that the overall rating was sufficient as a measure of SA. Unlike the SAGAT variables, all four SA ratings strongly correlated with participants' performance ratings and moderately correlated with the expert observer's performance ratings. The four subjective measures of SA slightly correlated with the workload measures. Although many of the correlations did not reach statistical significance, the general trend was as workload decreased SA increased. This relationship seems reasonable since periods of low workload provide an opportunity for more scanning and strip marking - actions which should lead to better SA.

A final question on the Post-Scenario Questionnaire requested that controllers rate the difficulty of the scenario they had just completed. A correlation of this rating with the other dependent variables in the experiment was used to determine what factors may have contributed to the difficulty of a scenario. The results indicated that difficulty ratings correlated with several performance measures including the number of altitude assignments, heading assignments, ground-to-air transmissions, aircraft density, and flights completed. Also, very strong correlations were found with the workload measures. Although the relationship was rather weak, the difficulty ratings were negatively correlated with the SA ratings. Finally, difficulty ratings were negatively correlated with the participants' and expert observer's performance ratings.

Table 8 shows the results of the correlational analysis conducted on the expert observer's ratings. One issue that was examined concerned the agreement between the expert's and participants' ratings. The results indicated a strong correlation between the expert's workload ratings and the participants' workload ratings and a weaker but significant correlation between the expert's performance ratings and the participants' performance ratings. These results suggest that there was agreement between the expert and participants and that they may have used similar indicators to make their ratings. An analysis of the expert's workload rating with the objective performance measures indicated significant correlations with the number of altitude assignments, heading assignments, ground-to-air transmissions, flights completed, and the aircraft density. Since these five taskload measures correlated with the participants' workload ratings as well, they may be the indicators that the expert and participants used to rate workload. The analysis also indicated that the expert's performance ratings negatively correlated with the number of conflicts and aircraft density. However, the participants' performance ratings did not correlate with either of these two performance measures. Also, it should be noted that the expert's ratings of workload and performance showed a strong negative correlation and were in agreement with the negative correlation found between the participants' ratings of workload and performance.

4.1.1 Summary and Conclusions.

One of the interests of the present research was to develop and evaluate an objective measurement technique for controller SA. An important finding from the correlational analysis was that the SAGAT variables did not correlate with participants' ratings of SA. This suggests that what controllers considered to be "SA" was different from the air traffic information requested by the SAGAT. Several participants commented that they were unable to answer many of the SAGAT questions, but they still felt they had good SA for information important to safe ATC. These comments may have been motivated by a desire not to "look bad" when they were unable to remember relevant information. On the other hand, the comments may indicate that remembering the SAGAT information is not necessary for good ATC performance. The fact that the SAGAT variables did not correlate with any of the objective or subjective measures of performance suggests that the latter alternative may be correct. Other controller comments and evaluations concerning the SAGAT will be discussed in the final results section.

The results indicated that participants' ratings of SA were positively correlated with the participants' and expert's ratings of performance. That is, when SA was high, performance was good, and when SA was low, performance was worse. This finding suggests that SA may be a contributing factor to performance and supports the claim that memory is an important part of ATC. Also, the results indicated that controller workload ratings were negatively correlated with the participants' and the expert's ratings of performance. That is, when workload was low,

TABLE 8. CORRELATION VALUES FOR THE EXPERT OBSERVER VARIABLES

	OBSWRK	OBSPFM
NCNF		-0.52*
NHOMISS		
NALT	0.63**	
NHDG	0.54*	
NSPD		
NPTT	0.62*	
CMAV	0.87**	-0.56*
NCOMP	0.67**	
SGTIR		
SGTSL		
ATWIT	0.85**	-0.54*
PSQWRK	0.85**	-0.57*
PSQPFM	-0.57*	0.51*
PSQSAW	-0.59*	0.54*
PSQCAL		0.56*
PSQPAL		
PSQPSV		
PSQDIF	0.85**	-0.54*
OBSWRK		-0.72**
OBSPFM	-0.72**	

indicates a significant effect, $p < 0.05$
indicates a significant effect, $p < 0.01$

perceived performance was good and when workload was high, perceived performance was worse. This finding suggests that controller workload may affect ATC performance. Controller workload ratings were negatively (but weakly) correlated with the participants' ratings of SA. That is, when workload was low, SA was generally high and when workload was high, SA was generally low. This finding suggests that controller workload may influence SA. However, it must be emphasized that these relationships between performance, SA, and workload are based upon the participants' and expert's subjective perceptions. These relationships were not confirmed by SAGAT or any of the objective performance measures.

4.2 MEMORY AIDS AND TRAFFIC VOLUME EFFECTS.

The results of the study concerning the effectiveness of the proposed memory aids in low and high traffic scenarios are reported in this section. The primary statistical procedure used will be ANOVA. In simple terms, ANOVA is a method for comparing the variability produced by the treatment (e.g., the memory aids) to the variability produced by other extraneous factors (e.g., individual controller style). The results of the analysis produce an F-ratio, and the larger the F-ratio the greater the effects of the treatment. An F-ratio is considered to be statistically significant if it exceeds a given critical value determined by the degrees of freedom in the experimental design. The analyses associated with each independent variable are referred to as main effects, and the analyses associated with combinations of variables are referred to as interaction effects.

An interaction occurs when the effects of one variable are different, depending upon the level of another variable. To examine the pattern of interactions between variables, the experimental design is often broken down into its basic components, referred to as simple main effects. In the present experiment, there are four simple main effects and each consists of a comparison between two means. One simple main effect involves the difference between the no-memory aids and memory aids conditions in low traffic scenarios and the second involves the differences between the memory aids conditions in high traffic scenarios. The third simple main effect involves the difference between low and high traffic conditions in scenarios without memory aids and the last involves the difference between the traffic conditions in scenarios with memory aids. If an interaction between variables is significant, an F-ratio is computed for each simple main effect to determine if the difference between the two means is statistically significant.

In this section, the approach will be to present graphical plots of the means (or frequency plots) for the main experimental conditions as a clear and simple summary of the results for each dependent variable. The important findings of the experiment will be reported through a discussion of the main trends in each graph. However, graphical plots can be misleading, since the variability in the means is not shown. Therefore, ANOVA will be used to confirm (or disconfirm) the apparent trends in each graph. Although all 20 dependent variables were analyzed, the results of only 13 variables will be presented in this summary. All the performance variables will be reported, except the number of speed assignments which was not an important variable in this study. Since ATWIT was highly correlated with the participants' and expert observer's post-scenario workload ratings, only the ATWIT results will be discussed. Also, since the participants' overall ratings of SA were highly correlated with the other subjective measures of SA, only the overall ratings will be presented. Finally, the participants' scenario difficulty ratings will not be reported. Tables of means and ANOVAs for all 20 dependent variables are provided in appendix C and appendix D, respectively.

Figure 1 presents the main results for the number of conflicts. Since the means were very small, a frequency plot of the number of conflict errors for the entire study is a more meaningful display of the data. Although a relatively rare event in the simulation, 28 conflicts were recorded in the

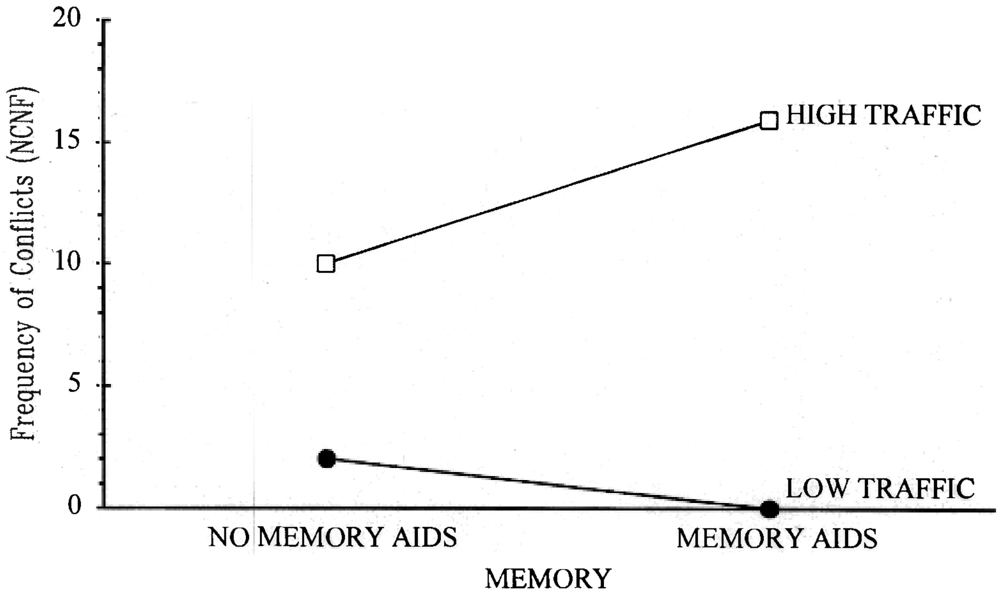


FIGURE 1. FREQUENCY OF CONFLICTS FOR THE ENTIRE EXPERIMENT AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

128 scenarios performed. Since an average of 18.5 aircraft were scheduled to appear in each 30-minute scenario, nearly 12 aircraft were involved in a separation error (with another aircraft) for every 1000 aircraft worked. In other terms, one conflict occurred almost every 2.3 hours.

The intentionally demanding conditions of the simulation produced far more conflict errors than actually occur in reality. As shown in the graph, almost all of the conflicts occurred during high traffic scenarios. Also, there seems to be slightly more conflicts in high traffic scenarios with the memory aids than without the memory aids. However, this difference is small and may be due to chance alone.

The results of the ANOVA indicated that the main effect of MEMORY, [$F(1,15) = 0.71$], was not significant. However, the main effect of TRAFFIC, [$F(1,15) = 12.17$, $p<0.01$], was significant. The MEMORY*TRAFFIC, [$F(1,15) = 2.73$], interaction was not significant. These statistical results indicated that the memory aids did not decrease (or increase) the incidence of conflict errors. The apparent trend for a greater number of conflicts in high traffic scenarios with the memory aids was not statistically reliable. Also, the analysis confirmed that more conflict errors occurred in high, relative to low, traffic scenarios.

Figure 2 displays the main results for the number of handoff misses. Again, since the means were very small, a frequency plot of the number of handoff errors for the entire study is a more meaningful display of the data. Handoff errors were even more rare than conflicts and 8 handoff misses were recorded in the 128 scenarios performed. Over 3 aircraft were not properly handed off for every 1000 aircraft worked, or one handoff error occurred every 8 hours. Fortunately, the

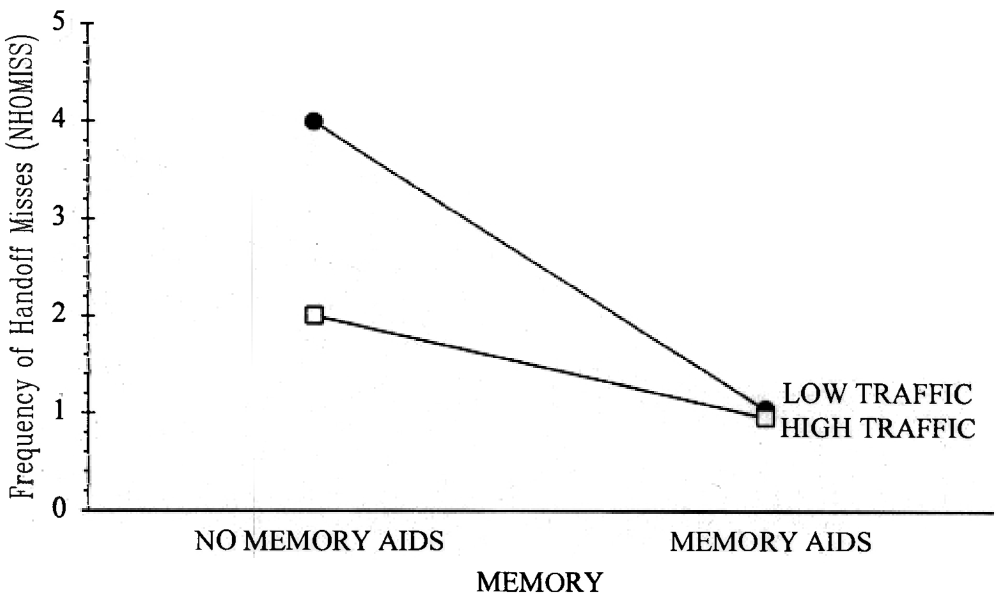


FIGURE 2. FREQUENCY OF HANDOFF MISSES FOR THE ENTIRE EXPERIMENT AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

incidence of handoff errors is much less frequent in reality. As shown in the graph, most of the handoff misses occurred in scenarios without the memory aids. Also, nearly the same number of handoff errors occurred in low and high traffic scenarios.

The results of the ANOVA revealed a significant main effect of MEMORY, [$F(1,15) = 5.00$, $p<0.05$]. The main effect of TRAFFIC, [$F(1,15) = 1.00$], and the interaction, MEMORY*TRAFFIC, [$F(1,15) = 1.23$], were not significant. These statistical results confirmed that the memory aids reduced the number of handoff errors. Also, the analysis indicated that there was no difference in handoff misses between low and high traffic scenarios.

The main results of the experiment for the number of altitude assignments heading assignments are summarized in figures 3 and 4. As shown, the general trends are very similar for these two variables. As expected, the number of altitude and heading assignments was greater in scenarios without the memory aids and high traffic scenarios. However, the effects of the memory aids appear to be slightly greater for high, relative to low, traffic scenarios. In other words, although the memory aids reduced the number of altitude and heading assignments in low traffic scenarios, the reductions were generally greater in high traffic scenarios.

The ANOVA revealed a significant main effect of MEMORY, [$F(1,15) = 369.45$, $p<0.01$, and $F(1,15) = 468.87$, $p<0.01$] for both types of assignments. Also, the main effect of TRAFFIC,

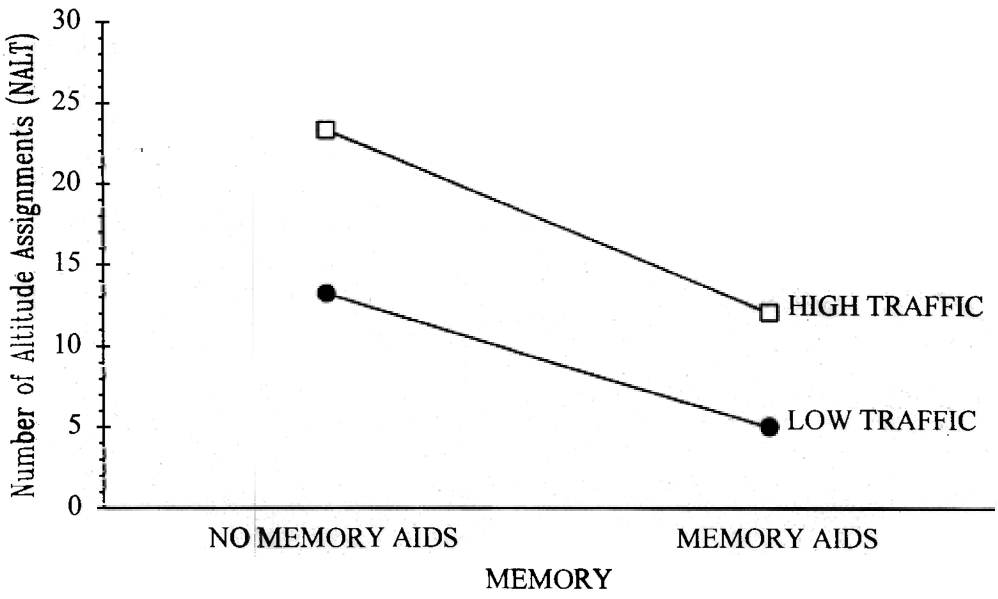


FIGURE 3. MEAN NUMBER OF ALTITUDE ASSIGNMENTS AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

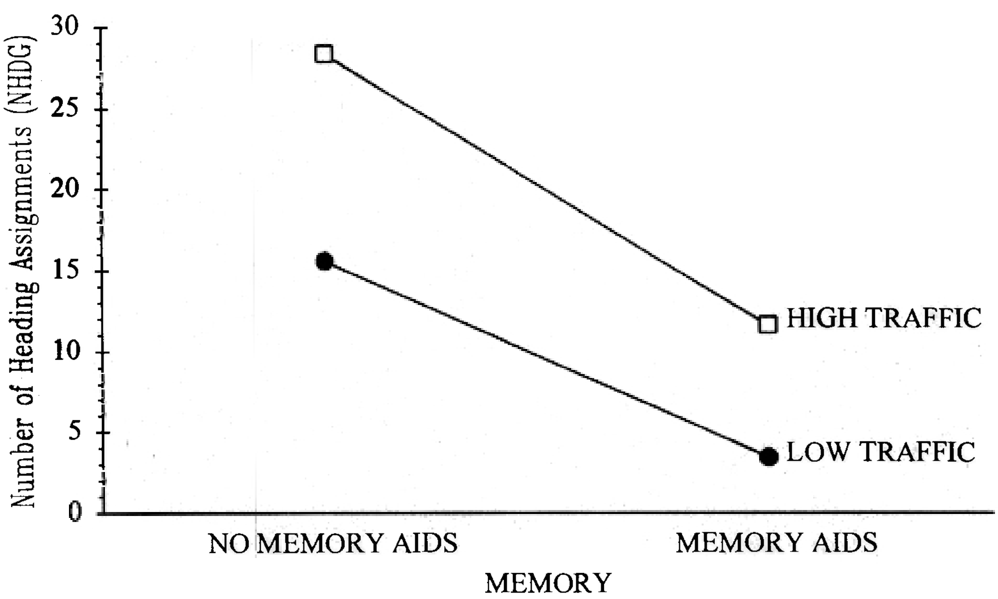


FIGURE 4. MEAN NUMBER OF HEADING ASSIGNMENTS AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

[F(1,15) = 457.59, p<0.01, and F(1,15) = 375.22, p<0.01], was significant for both assignments. However, these main effects must be qualified because the MEMORY*TRAFFIC, [F(1,15) = 12.31, p<0.01, and F(1,15) = 55.89, p<0.01], interaction was significant for both types of

assignments. Since the interactions were significant, an analysis of simple main effects was conducted on each variable. A summary of the analysis for the number of altitude and heading assignments is presented in tables 9 and 10. As shown, all four simple main effects were significant for both types of assignments. The interaction occurred because the difference between the no memory aids and memory aids conditions was greater for high traffic scenarios.

The main results of the experiment for the number of ground-to-air transmissions are reported in figure 5. As shown, the general trends are similar to the number of altitude and heading assignments in the previous graphs. The number of ground-to-air transmissions was greater for scenarios without the memory aids and high traffic scenarios. However, the memory aids appear to have similar effects in low and high traffic scenarios.

The ANOVA revealed significant main effects of both MEMORY, [F(1,15) = 26.36, p<0.01], and TRAFFIC, [F(1,15) = 1131.89, p<0.01]. The MEMORY*TRAFFIC, [F(1,15) = 0.67], interaction was not significant. These statistical results confirmed that the memory aids reduced the number of ground-to-air transmissions in all scenarios. Also, the analysis indicated that there were more ground-to-air transmissions in high relative to low traffic scenarios.

TABLE 9. ANALYSIS OF SIMPLE MAIN EFFECTS FOR THE NUMBER OF ALTITUDE ASSIGNMENTS

TRAFFIC	NO MEMORY AIDS	MEMORY AIDS	Difference	F-Ratio
LOW	13.25	5.03	8.22	267.51**
HIGH	23.34	12.13	11.21	201.38**

MEMORY	LOW TRAFFIC	HIGH TRAFFIC	Difference	F-Ratio
NO MEMORY AIDS	13.25	23.34	-10.09	232.88**
MEMORY AIDS	5.03	12.13	-7.10	201.29**

-----
* indicates a significant effect, p < 0.05
** indicates a significant effect, p < 0.01

TABLE 10. ANALYSIS OF SIMPLE MAIN EFFECTS FOR THE NUMBER OF HEADING ASSIGNMENTS

TRAFFIC	NO MEMORY AIDS	MEMORY AIDS	Difference	F-Ratio
LOW	15.59	3.47	12.12	363.00**
HIGH	28.38	11.66	16.72	416.41**

MEMORY	LOW TRAFFIC	HIGH TRAFFIC	Difference	F-Ratio
NO MEMORY AIDS	15.59	28.38	-12.79	169.71**
MEMORY AIDS	3.47	11.66	-8.19	431.05**

-----

* indicates a significant effect, $p < 0.05$

** indicates a significant effect, $p < 0.01$

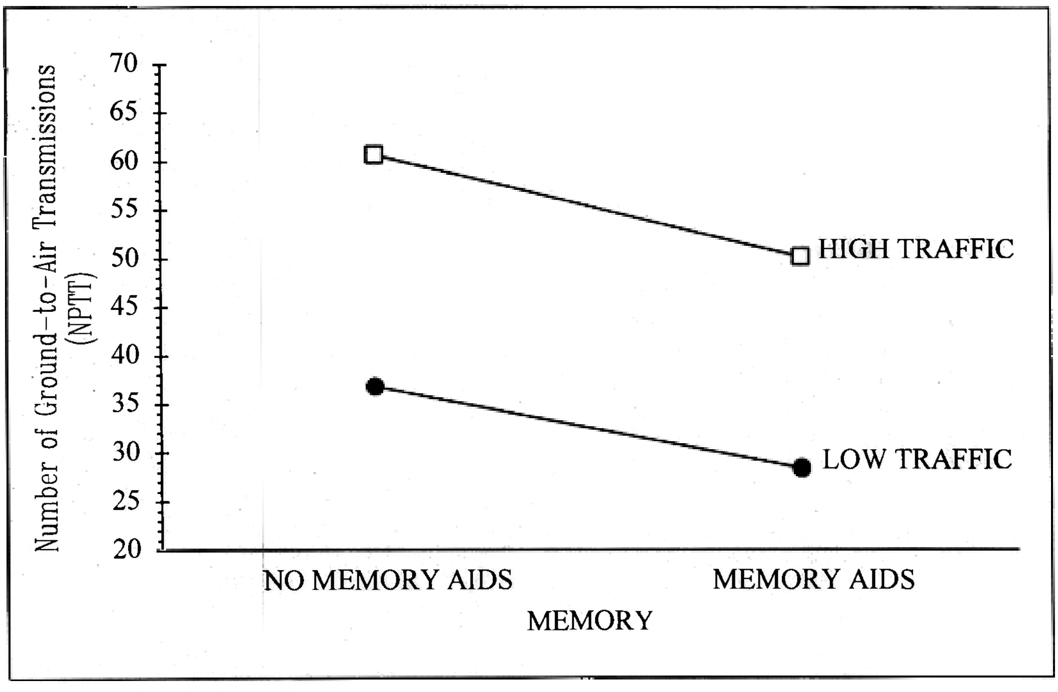


FIGURE 5. MEAN NUMBER OF GROUND-TO-AIR TRANSMISSIONS AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

Figure 6 presents the main results for the cumulative average of system activity or aircraft density. As planned in the design of scenarios, the aircraft density for high traffic scenarios was greater than for low traffic scenarios. Also, aircraft density was almost identical for scenarios with memory aids and without memory aids.

The results of the ANOVA indicated that the main effect of MEMORY, [$F(1,15) = 1.31$], was not significant. However, the main effect of TRAFFIC, [$F(1,15) = 2265.52$, $p<0.01$], was significant. The MEMORY*TRAFFIC, [$F(1,15) = 0.01$], interaction was not significant. These statistical results indicated that the memory aids had no effect on aircraft density. Also, the analysis confirmed that aircraft density was greater in high, relative to low, traffic scenarios.

Figure 7 displays the main results for the number of flights completed. As shown, the trends are very similar to the aircraft density graph. As expected, the number of flights completed for high traffic scenarios was greater than for low traffic scenarios. Also, there seems to be little difference in the number of flights completed between scenarios with memory aids and without memory aids.

The results of the ANOVA indicated that the main effect of MEMORY, [$F(1,15) = 1.17$], was not significant. However, the main effect of TRAFFIC, [$F(1,15) = 603.39$, $p<0.01$], was significant. The MEMORY*TRAFFIC, [$F(1,15) = 0.29$], interaction was not significant. These statistical results indicated that the memory aids did not affect the number of flights completed. Also, the analysis confirmed that the number of flights completed was greater for high, relative to low, traffic scenarios.

The main results of the study for the controller's performance rating and the expert observer's performance rating are summarized in figures 8 and 9. As shown, there is an amazing degree of similarity in both the magnitude and pattern of these two ratings. First, both graphs indicate that performance ratings were higher in low traffic scenarios. Also, there is a slight diverging pattern in both graphs. That is, performance ratings seem to increase in low traffic scenarios, but decrease in high traffic scenarios when using the memory aids. However, this interesting tendency is rather small and may be due to chance alone.

The ANOVA indicated that the main effect of MEMORY, [$F(1,15) = 0.13$, and $F(1,15) = 0.13$], was not significant for either set of ratings. However, the main effect of TRAFFIC, [$F(1,15) = 28.01$, $p<0.01$, and $F(1,15) = 72.61$, $p<0.01$], was significant for both ratings. The MEMORY*TRAFFIC, [$F(1,15) = 4.01$, and $F(1,15) = 4.41$], interaction was not significant for either set of ratings. These statistical results indicated that the memory aids did not increase (or decrease) either of the performance ratings. The apparent trend for the memory aids to increase ratings in low traffic scenarios and decrease ratings in high traffic scenarios was not statistically reliable. Also, the analysis confirmed that both performance ratings were higher in low, relative to high, traffic scenarios.

The main results of SAGAT for the information recall phase and the spatial location phase are reported in figures 10 and 11. As shown, there is a large degree of similarity in both of these objective measures of SA. First, both graphs indicate that SA was higher in low

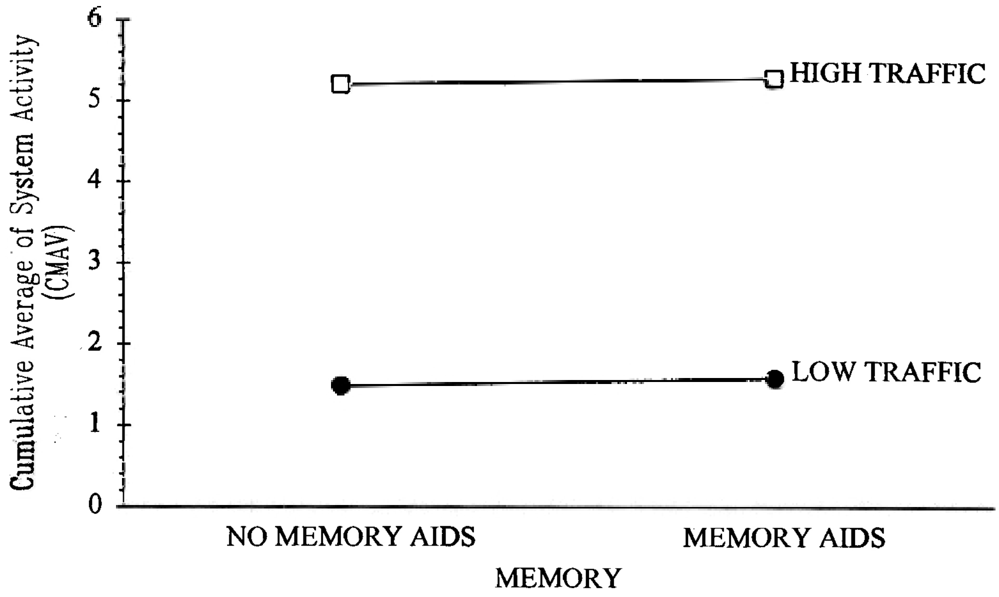


FIGURE 6. MEAN CUMULATIVE AVERAGE OF SYSTEM ACTIVITY AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

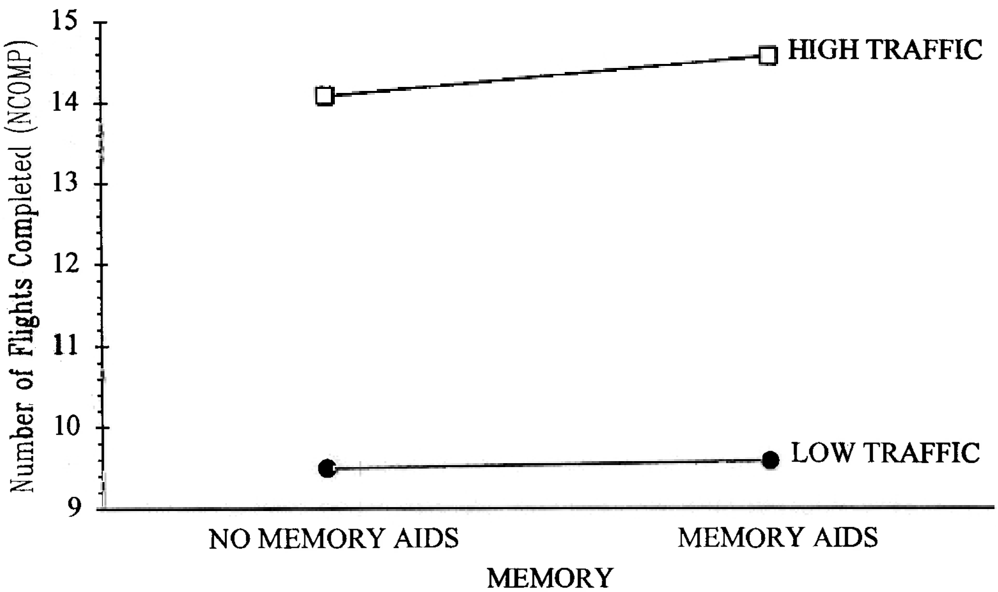


FIGURE 7. MEAN NUMBER OF FLIGHTS COMPLETED AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

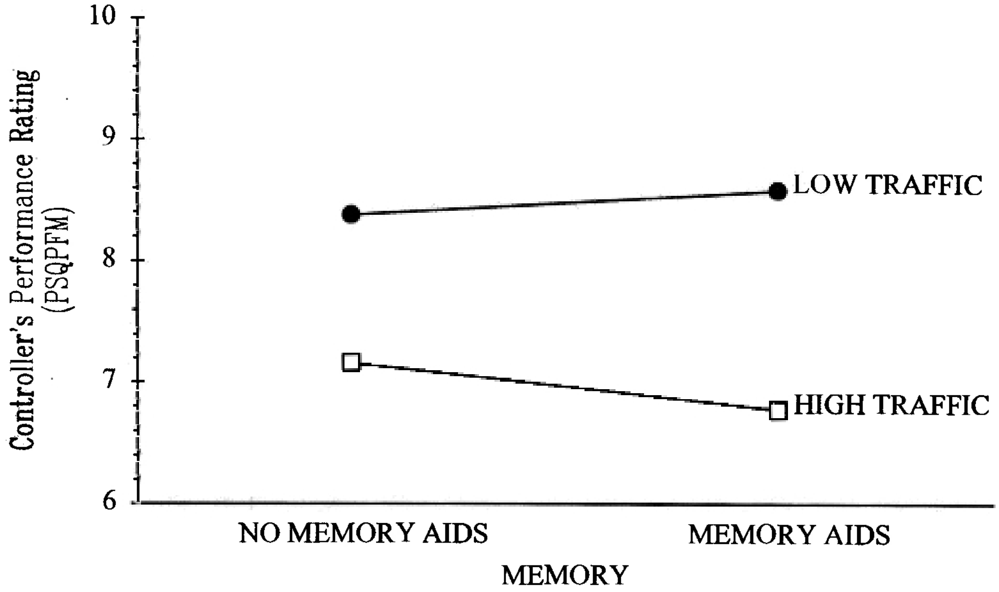


FIGURE 8. MEAN CONTROLLER'S PERFORMANCE RATING AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

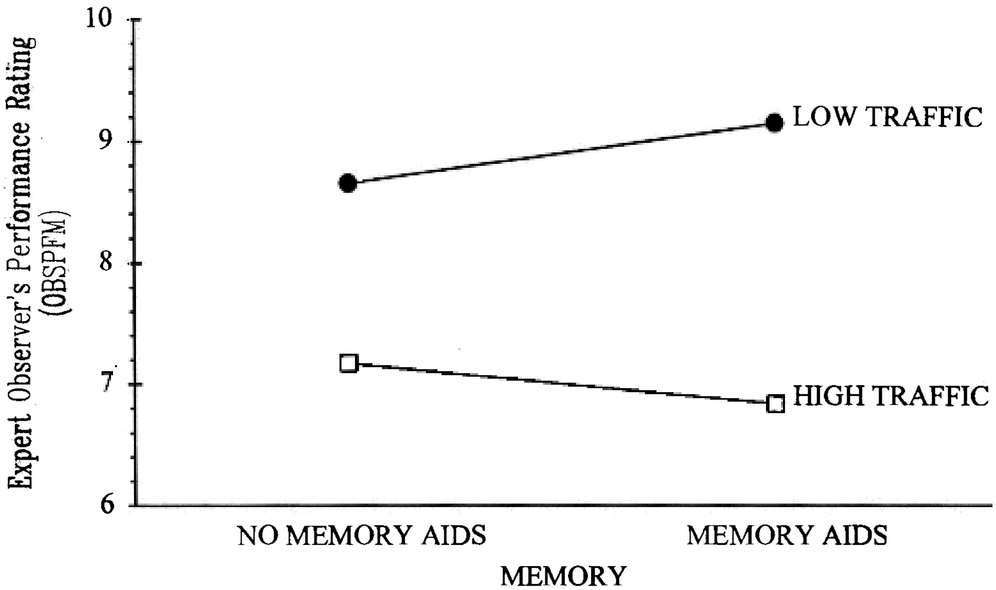


FIGURE 9. MEAN EXPERT OBSERVER'S PERFORMANCE RATING AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

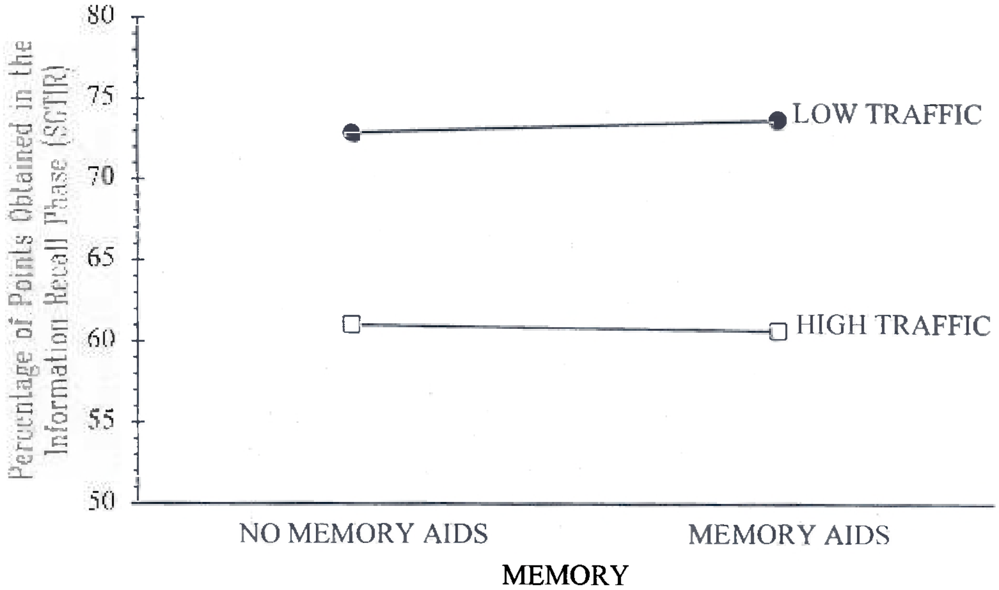


FIGURE 10. MEAN PERCENTAGE OF POINTS OBTAINED IN THE INFORMATION RECALL PHASE AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

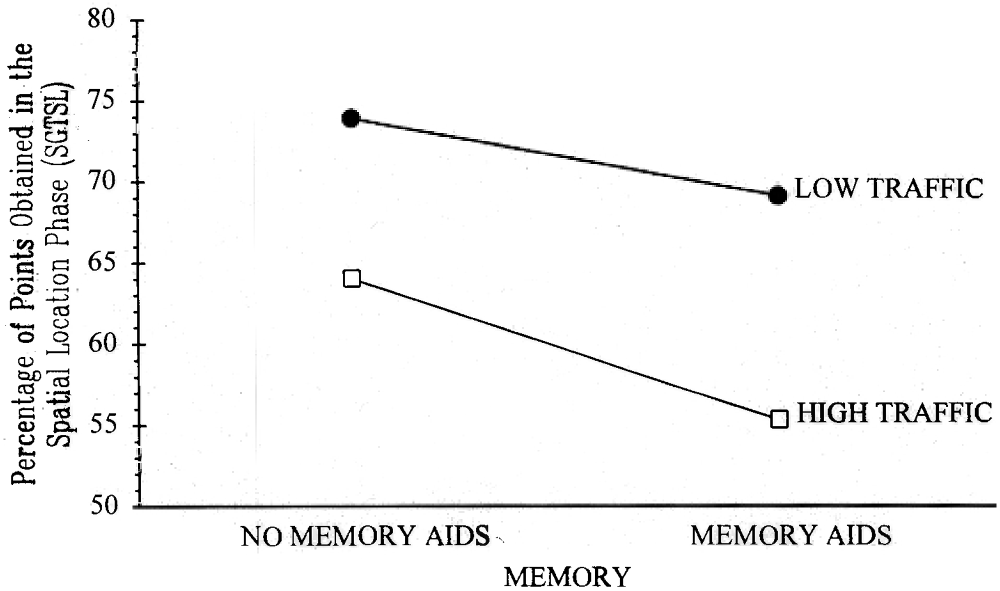


FIGURE 11. MEAN PERCENTAGE OF POINTS OBTAINED IN THE SPATIAL LOCATION PHASE AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

traffic scenarios. Contrary to expectations, there was little difference in SA between scenarios with memory aids and without memory.

The results of the ANOVA indicated that the main effect of MEMORY, [$F(1,15) = 0.01$, and $F(1,15) = 2.65$], was not significant for either objective measure. However, the main effect of TRAFFIC, [$F(1,15) = 10.99$, $p<0.01$, and $F(1,15) = 19.95$, $p<0.01$], was significant for both measures. The MEMORY*TRAFFIC, [$F(1,15) = 0.03$, and $F(1,15) = 0.33$], interaction was not significant for either objective measure. These statistical results indicated that the memory aids did not affect SA in either the information recall or spatial location phases. Also, the analysis confirmed that SA was higher in low relative to high traffic scenarios for both phases.

The main results of the experiment for the controller's overall SA rating are summarized in figure 12. As shown, the participants' ratings of SA were higher in low traffic scenarios. This trend is in agreement with both SAGAT measures of SA. In contrast, the graph indicates that SA ratings may be slightly higher in scenarios with the memory aids. However, this tendency is rather small and may be due to chance alone.

The ANOVA indicated a non-significant main effect for MEMORY, [$F(1,15) = 3.46$], and a significant effect for TRAFFIC, [$F(1,15) = 31.29$, $p<0.01$]. The MEMORY*TRAFFIC, [$F(1,15) = 0.27$], interaction was not significant. These statistical results confirmed that the memory aids may have slightly increased SA ratings (the effect would be significant at $p<.10$). Although this result was not strong enough to reach statistical significance at a standard level ($p<0.05$), it is reported here because of its importance to the primary goal of this research. Also, the analysis indicated that SA ratings were higher in low, relative to high, traffic scenarios.

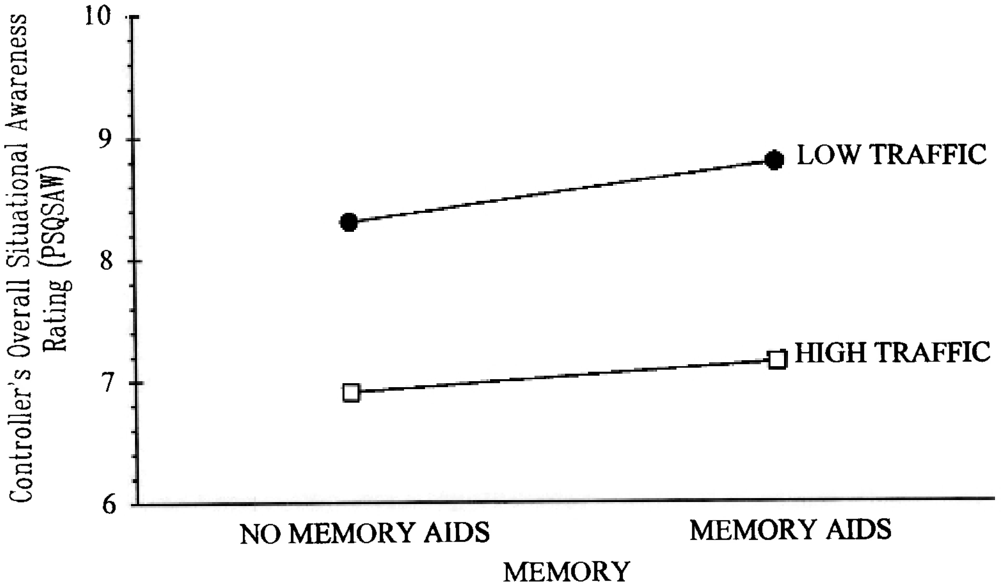


FIGURE 12. MEAN CONTROLLER'S OVERALL SITUATIONAL AWARENESS RATING AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

Figure 13 presents the main results for ATWIT ratings. As shown, controllers' workload ratings for high traffic scenarios were higher than for low traffic scenarios. Surprisingly, there was little difference in workload ratings between scenarios with memory aids and without memory aids. These trends are very similar to the patterns produced in the graphs of aircraft density and number of flights completed.

The results of the ANOVA indicated that the main effect of MEMORY, [F(1,15) = 0.01], was not significant. However, the main effect of TRAFFIC, [F(1,15) = 147.63, p<0.01], was significant. The MEMORY*TRAFFIC, [F(1,15) = 0.18], interaction was not significant. These statistical results indicated that the memory aids did not affect workload ratings. Also, the analysis confirmed that workload ratings were higher in high, relative to low, traffic scenarios.

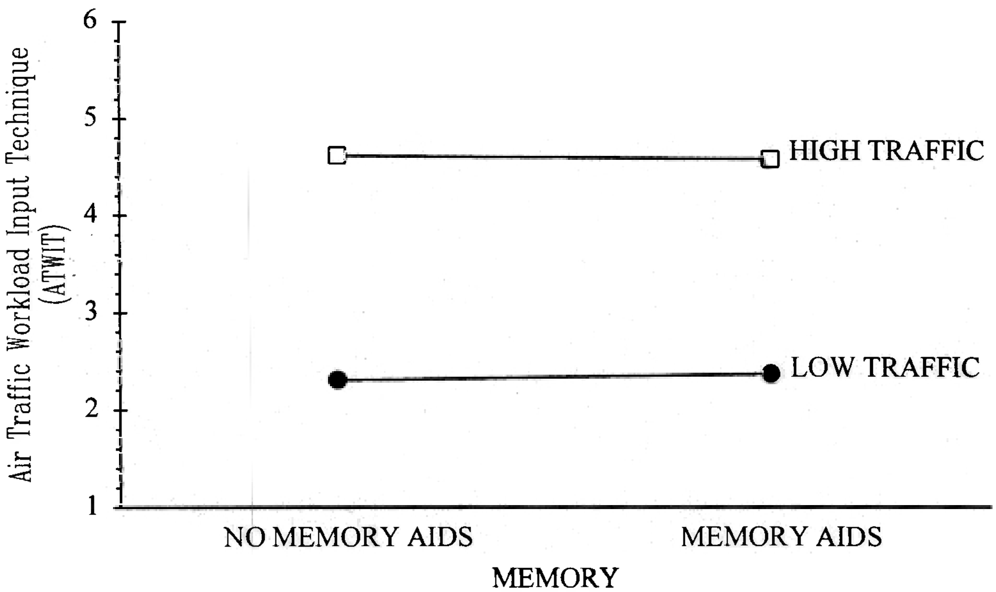


FIGURE 13. MEAN AIR TRAFFIC WORKLOAD INPUT TECHNIQUE AS A FUNCTION OF THE MEMORY AIDS AND TRAFFIC IN SCENARIOS

A general summary of all the ANOVA results is provided in table 11. The table values indicate whether the main effects or interaction effect for each variable were statistically significant (YES) or not statistically significant (NO).

TABLE 11. A SUMMARY OF THE ANOVA RESULTS FOR EACH DEPENDENT VARIABLE IN THE EXPERIMENT

Variable	MEMORY EFFECT	TRAFFIC EFFECT	INTERACTION EFFECT
NCNF	NO*	YES**	NO
NHOMISS	YES	NO	NO
NALT	YES	YES	YES
NHDG	YES	YES	YES
NPTT	YES	YES	NO
CMAV	NO	YES	NO
NCOMP	NO	YES	NO
PSQPFM	NO	YES	NO
OBSPFM	NO	YES	NO
SGTIR	NO	YES	NO
SGTSL	NO	YES	NO
PSQSAW	NO	YES	NO
ATWIT	NO	YES	NO

* NO - NOT STATISTICALLY SIGNIFICANT
** YES - STATISTICALLY SIGNIFICANT

4.2.1 Summary and Conclusions.

One of the purposes of the memory aids was to reduce the need for communications between controllers and pilots. It was reasoned that fewer communications would allow controllers more time to scan the radar display, review flight progress strips, and perform other activities that should increase their SA. The results of the experiment confirmed that the memory aids greatly decreased the number of ground-to-air transmissions, including both altitude and heading assignments. The memory aids reduced the number of altitude and heading assignments in both low and high traffic scenarios, but the reductions (and potential benefits) were greater in high traffic scenarios. The reduction in communications is an important finding since many misunderstandings and errors occur in ATC communications.

The results regarding the potential benefits of the memory aids for improving other aspects of performance were mixed. Although conflicts and handoff misses are major concerns in ATC, the frequency of these errors is rare. Therefore, any conclusions based upon the few incidents that did occur during the study must be considered as tentative. Contrary to expectations, the results indicated that the number of conflict errors was nearly equal in scenarios with the memory aids and without the memory aids. Although a disproportionate number of conflicts occurred while using the memory aids in high traffic scenarios, this trend was not reliable. In contrast, there were slightly fewer handoff errors while using the memory aids. There was a strong consistency between the participants' and expert observer's ratings of performance. The ratings indicated that the memory aids slightly increased performance in low traffic scenarios, while there was a slight decrease in high traffic scenarios. Overall, these results suggest that the memory aids may slightly improve performance when traffic volume is relatively low. However, when traffic volume is extremely high, the memory aids may slightly detract from performance.

The results concerning the potential benefits of the memory aids for improving SA were mixed as well. Using SAGAT as an objective measurement technique, the memory aids did not increase SA. Contrary to expectations, the accuracy of controllers' responses in both the information recall and spatial location phases was not improved by the memory aids. However, using a subjective measurement technique, the memory aids did increase SA. Participants' ratings of SA improved slightly in scenarios with the memory aids. The results from the objective measures and subjective ratings were expected to agree and indicate improvements in SA using the memory aids. However, participants' ratings may be a more sensitive measure, while the present SAGAT format may be unable to detect small differences in SA. Perhaps the most reasonable conclusion regarding the effectiveness of the memory aids for improving SA is that the benefits are very small.

A secondary goal of this study was to investigate the effects of the memory aids on controller workload. Surprisingly, the results indicated that the memory aids had no effect on controller workload. Controllers' ATWIT ratings were nearly identical in scenarios with memory aids and without memory aids. There are at least two possible explanations for this unexpected result. First, although the memory aids greatly reduced communications, controllers may have felt that the communications were not contributing much to their workload in the first place. However, this explanation seems unlikely since most controllers would probably agree that communications are a major source of workload. A more reasonable explanation is that the memory aids had both beneficial effects (tending to decrease workload) and detrimental effects (tending to increase workload) that were nearly equal in strength. Most likely, the reduction in communications had a beneficial effect on workload. However, the memory aids consisted of arrival and departure procedures that were not routinely used by controllers. Perhaps, having to consider this relatively unfamiliar information while controlling traffic had a detrimental effect on workload as well. The net result of these beneficial and detrimental effects may have had a canceling effect that led to no change in controllers' workload ratings.

The results of the study indicated that traffic volume had strong effects on ATC performance, SA, and workload. For almost every dependent variable examined, there were large differences between low and high traffic scenarios (exception: number of handoff misses). As expected, there were many more conflicts, altitude assignments, heading assignments, and ground-to-air transmissions in high traffic scenarios. Also, it was not surprising that aircraft density was much greater and there were more flights completed in high traffic scenarios. Performance was lower in high traffic scenarios as indicated by the participants' and expert observer's ratings. SA was lower in high traffic scenarios as shown in the SAGAT variables and the participants' ratings.

Controller workload was higher in high traffic scenarios as revealed by the participants' ATWIT ratings.

4.3 FINAL CONTROLLER RATINGS AND COMMENTS.

The results of the study concerning controllers' opinions and comments about the experiment are reported in this section. Controllers' ratings obtained from the Final Questionnaire are summarized in tables that follow. In addition to providing ratings and comments regarding the effectiveness, usability, and acceptability of the experimental arrival and departure procedures, this questionnaire requested opinions regarding a third potential memory aid referred to as "completing the transaction." This communication technique required controllers to issue all possible assignments to aircraft on initial radar contact. As experimental testing began, it became evident that most controllers already used this technique routinely without special instructions. The usefulness of this communication technique was limited because the experimental arrival and departure procedures reduced the need for assignments on initial radar contact. However, completing the transaction can potentially benefit controllers who do not use it already by reducing communications and the need to remember planned assignments. A summary of controllers' ratings regarding this technique is presented, as well as ratings of the memory aids that were more important to the experiment. The summary includes several ratings concerning the realism of the simulation and the intrusiveness of the measurement techniques.

One of the problems with this final ratings technique is that there is no obvious standard of comparison. Therefore, one reasonable way to interpret the results is to compare the ratings to the mid-point (indifference point) of the scale. For most of the responses in this questionnaire, mean ratings which exceeded 5.5 on the 10-point scale can be thought of as indicating a favorable outcome. Although it may be tempting to compare these final ratings to the results of the post-scenario ratings, these measurement techniques are very different and the comparison would have little meaning. The experimental methods used to collect and analyze the post-scenario ratings were scientifically more valid. However, these final ratings provided an additional means for evaluating the effectiveness of the proposed memory aids. Also, the questions were designed to elicit further comments from controllers which could benefit the research.

A summary of controllers' ratings regarding the effectiveness of the proposed memory aids for improving performance, increasing SA, and reducing workload is shown in table 12. Although there was some disagreement, controllers ratings of effectiveness were generally high. The results indicated that controllers thought that the memory aids were most useful for reducing their workload, slightly less useful for improving performance, and the least useful for increasing SA. Although the difference was very small, departure procedures were rated as generally more effective than arrival procedures. Also, completing the transaction was thought to improve performance a great deal, while the effects on SA were relatively small. In addition to these

TABLE 12. A SUMMARY OF CONTROLLERS' RATINGS OF MEMORY AIDS EFFECTIVENESS

How useful were the memory aids for improving, increasing, and reducing the specified attributes?

1	2	3	4	5	6	7	8	9	10
Not Very Useful									Extremely Useful

Memory Aids	Performance Mean (SD)	Awareness Mean (SD)	Workload Mean (SD)
Arrival Procedures	7.00 (2.07)	6.38 (2.28)	7.50 (2.10)
Departure Procedures	7.13 (1.96)	6.69 (2.06)	7.88 (1.67)
Completing the Transaction	8.13 (1.45)	6.38 (2.63)	7.81 (1.87)

ratings, controllers' mentioned several important issues regarding the effectiveness of the memory aids in their comments. Many controllers thought that they could have done better in the simulation if they had been more familiar with the proposed memory aids. One controller commented, "The arrival and departure procedures would be much more useful if there was more time to study them and commit them to memory." It seems the training program that was developed to familiarize controllers with the memory aids was not totally successful. In retrospect, it may have been expecting too much from controllers to be able to memorize the experimental arrival and departure procedures completely within the 1-hour training session. Although a display of the procedures was placed nearby for controllers to review as they worked traffic, it is likely that the memory aids would have been more effective if this information was committed to memory. Another controller stated, "I felt that the arrival and departure procedures were helpful, but I worked harder mentally." In agreement with results presented earlier, comments like these suggest that the memory aids had mixed effects on controller workload. Although the arrival and departure procedures reduced communications and physical workload, the lack of familiarity with the memory aids may have increased mental workload.

Another issue mentioned by some controllers was that they were uncomfortable with the memory aids because they felt they lost control over the aircraft. One controller commented, "I think I was fooled into thinking I didn't have to do anything. I felt that I wasn't really in control of the situations, but that I was just reacting." A few controllers made comments like this even though they were instructed that the memory aids did not guarantee aircraft separation. The arrival and departure procedures can be thought of as a form of automation where certain functions were done for controllers without them needing to take any action. Although automation is more commonly thought of as being performed by computers, pseudo pilots accomplished the functions in the present simulation. There are always potential risks associated with automation in that the human operator may become more passive, less aware, and less able to respond when control actions are required. These negative consequences may have been experienced by a few controllers in the present experiment. As one controller stated, "Actually, I think the arrival and departure procedures decreased my SA because I wasn't actually controlling all the action." Although there are many potential benefits of automation, there are risks as well, and one important contribution of human factors research is to evaluate the impact of automation on controller effectiveness.

Several controllers raised some other concerns about potential disadvantages of the memory aids. The arrival procedures were designed to structure the flow of traffic so that aircraft arriving from different directions would follow the same flight path in the near vicinity of the airport. It was reasoned that it might benefit controllers to establish a standard approach for all aircraft, instead of having to assign different vectors to aircraft in order to assist their approach to the airport. One controller commented, "Sometimes you can save the pilot some time (and fuel) by taking him off the arrival and giving him a direct vector." Another controller stated, "The arrival procedures increased workload by putting aircraft in proximity to aircraft that wouldn't normally happen with vectors." For the arrival procedures used in the present study, both of these comments seem to have some merit and should be taken into consideration. If the arrival procedures were designed differently, it may have been possible to minimize these potential disadvantages.

A summary of controllers' ratings regarding the usability and acceptability of the proposed memory aids is shown in tables 13 and 14. Usability refers to how easy the memory aids were to use, while acceptability refers to controllers' willingness to use the memory aids. Although these issues may seem less important than effectiveness, they are nonetheless relevant to the final evaluation of the proposed memory aids. For example, if the memory aids were invoked by an awkward set of control instructions to pilots or a lengthy series of keystrokes on the keypack, then usability would be poor. Also, if controllers were not confident that the memory aids would benefit them or function reliably, then acceptability would be low. Although no usability problems were expected for the memory aids employed in the present study, acceptability was less certain. As shown, controllers' ratings of both usability and acceptability were very high. However, a few controllers expressed some apprehension about using the arrival procedures stating that they just felt safer by taking the aircraft off the arrivals.

Table 15 presents a summary of controllers' ratings regarding the realism of the simulation. In general, the controllers were very impressed with the laboratory and agreed that the realistic equipment and testing conditions created an extremely high-fidelity simulation. Most of the controllers stated that the radar display of ACY airspace was very accurate. One controller commented, "The only portion of the simulation that differed a great deal was the altitude structure and this was not hard to get used to at all." Although some minor problems were noted, most controllers thought the traffic scenarios were realistic. Two controllers mentioned, "There was much more IFR traffic than we're used to," and "Most moderate or extremely busy days are

TABLE 13. A SUMMARY OF CONTROLLERS' RATINGS OF MEMORY AIDS USABILITY

Please rate how easy the memory aids were to use.

1	2	3	4	5	6	7	8	9	10
Not Very Easy									Extremely Easy

Memory Aids	Usability Mean (SD)
Arrival Procedures	8.19 (1.80)
Departure Procedures	8.31 (1.70)
Completing the Transaction	8.38 (1.54)

TABLE 14. A SUMMARY OF CONTROLLERS' RATINGS OF MEMORY AIDS ACCEPTABILITY

Please rate your willingness to use the memory aids.

1	2	3	4	5	6	7	8	9	10
Not Very Willing									Extremely Willing

Memory Aids	Acceptability Mean (SD)
Arrival Procedures	7.81 (1.97)
Departure Procedures	8.06 (1.98)
Completing the Transaction	8.63 (1.31)

TABLE 15. A SUMMARY OF CONTROLLERS' RATINGS OF SIMULATION REALISM

How realistic were the specified characteristics of the simulation?

1	2	3	4	5	6	7	8	9	10

Not Very Realistic — Extremely Realistic

Memory Aids	Realism Mean (SD)
Overall Rating	7.88 (0.89)
Atlantic City Airspace	8.69 (0.95)
Low Traffic Scenarios	6.75 (1.53)
High Traffic Scenarios	7.06 (1.24)

almost always VFR conditions." Another controller stated, "The sector would probably be split up during the heavy sessions." These comments will be considered in future research. However, some slight deviations from reality were necessary in the present experiment. The memory aids were intended for IFR traffic and it was of interest to study their effects under both normal and unusually busy conditions. Therefore, IFR scenarios were constructed and controllers were asked to work the entire sector alone.

Table 16 displays a summary of controllers' ratings regarding the intrusiveness of the measurement techniques. Although ATWIT and SAGAT were not expected to interfere with the participants' ATC duties, data were collected from controllers to confirm this expectation. As shown, controllers' ratings were extremely low indicating that neither technique interfered very much with performance. Most controllers commented that ATWIT was very easy to use and that SAGAT was not a major distraction for them. Therefore, these results support the low interference concept of ATWIT and SAGAT for use in ATC simulation research. However, it became clear from controllers comments that SAGAT could be improved.

A criticism made by several participants was that it is not necessary for controllers to remember aircraft call signs to perform their job. In fact, the present SAGAT procedure greatly depended on controllers' memory for aircraft call signs. In the information recall and spatial locations phases, controllers were required to remember specific information (such as altitude and heading or the location on the radar display) about a particular aircraft that was denoted by its call sign. Many controllers stated that they could remember the locations of many of the aircraft on their

TABLE 16. A SUMMARY OF CONTROLLERS' RATINGS OF INTERFERENCE FROM THE MEASUREMENT TECHNIQUES

To what extent did the specified measurement techniques interfere with your performance?

1	2	3	4	5	6	7	8	9	10
Not Very Much									A Great Deal

Measurement Techniques	Interference Mean (SD)
ATWIT	1.75 (1.06)
SAGAT	2.63 (1.67)

radar display, but not the call signs of the aircraft. In other words, controllers thought of the aircraft as targets that needed to be separated and directed according to their flight plans, but specific identification of the targets by call sign was not necessary except for communications. These comments provided valuable insight into controller SA that will be considered when using SAGAT in future ATC simulations.

Controllers' ratings obtained from the Information Importance Form are summarized in table 17. In this table, one important result to note is the aircraft information that controllers considered to be the most important to their job. As shown, the highest importance ratings (on both days) were given to current altitude, most recently assigned altitude, current location, and arrival airport. Another important result to note is the agreement among controllers as to the importance of the aircraft information. Controllers generally agreed that current altitude, most recently assigned altitude, current location, and arrival airport were important information as indicated by the low standard deviations. Although the importance ratings for aircraft call sign were rather high on the first day, the ratings were much lower and more variable on the second day. The inconsistency in the ratings suggests that controllers did not agree on the importance of aircraft call signs. Lastly, the correlation between importance ratings on the two days is reported, but is a less relevant measure of consistency.

4.3.1 Summary and Conclusions.

Controllers' final ratings were generally favorable regarding the effectiveness of the proposed memory aids for improving performance, increasing SA, and reducing workload. However, many controllers thought that the memory aids would have been more useful if there had been more time to become familiar with the new procedures. Some controllers expressed concern about using the memory aids because they felt like they were no longer in control of aircraft.

TABLE 17. A SUMMARY OF CONTROLLERS' RATINGS OF IMPORTANCE FOR AIRCRAFT INFORMATION

Please rate the importance of the following aircraft information for controller performance.

1	2	3	4	5	6	7	8	9	10
Not Very Important							Extremely Important		

Type of Information	Day 1 Mean (SD)	Day 2 Mean (SD)	Correlation Day 1/Day 2
Aircraft Call Sign	9.38 (1.36)	7.19 (3.08)	0.11
Aircraft Type	8.06 (1.61)	6.88 (2.63)	0.76**
Aircraft Beacon Code	5.69 (3.11)	4.44 (2.45)	0.57*
Controller Ownership	7.25 (3.09)	7.06 (3.19)	0.84**
Entry Altitude	7.75 (2.24)	7.19 (2.59)	0.76**
Entry Airspeed	5.94 (2.17)	4.44 (2.28)	0.22
Entry Fix	7.19 (2.34)	6.88 (2.83)	0.65*
Exit Altitude	7.25 (2.79)	7.38 (2.50)	0.51**
Exit Airspeed	5.06 (1.91)	3.69 (2.12)	0.45
Exit Fix	7.06 (1.95)	6.75 (2.67)	0.40
Arrival Airport (within Sector)	8.81 (1.11)	8.06 (1.95)	0.68**
Departure Airport (within Sector)	8.13 (1.75)	7.75 (2.02)	0.58*
Current Altitude	9.06 (1.12)	8.94 (1.06)	0.12
Current Airspeed	6.25 (2.18)	5.75 (2.32)	0.50*
Current Heading	7.88 (1.89)	7.31 (2.39)	0.51*
Current Aircraft Location	9.13 (1.20)	8.56 (1.26)	0.48
Current Bad Weather Location	7.00 (1.71)	6.75 (2.49)	0.39
Most Recently Assigned Altitude	9.06 (1.44)	8.75 (1.18)	0.52*
Most Recently Assigned Airspeed	7.25 (2.05)	6.63 (2.16)	-0.19
Most Recently Assigned Heading	8.06 (1.65)	7.56 (2.13)	0.75**
Aircraft Holding/Spinning	7.19 (2.48)	6.75 (2.62)	0.70**
Aircraft Waiting for Handoff/Release	7.31 (2.27)	6.50 (2.50)	0.72**
Aircraft Near Exit Fix/Arrival Airport	8.31 (1.25)	8.19 (1.17)	0.55*
Density of Aircraft on Radar Display	7.31 (2.33)	8.00 (1.75)	0.31

* indicates a significant effect, $p < 0.05$
** indicates a significant effect, $p < 0.01$

Also, a few controllers mentioned that there were other disadvantages in using the arrival procedures relative to assigning direct vectors, such as increased flight time and fuel consumption. Finally, controllers' ratings were very favorable concerning the usability and acceptability of the memory aids.

Controllers' ratings of simulation realism were very high, and there were many positive comments about the accuracy of the airspace and traffic scenarios. Also, the experimenters observed that the participant controllers were very motivated to do well and took the simulation very seriously. Given the fidelity of the simulation and motivation of controllers, the results of the present study should be directly applicable to "real-world" ATC. Controllers generally agreed that the measurement techniques that were used during the simulation did not interfere with their efforts to control aircraft. Several comments suggested how the present SAGAT procedure could be improved in future research. Controllers' ratings indicated that the aircraft information that is most important for their work is current altitude, most recently assigned altitude, current location, and arrival airport. This information will be considered when planning future memory enhancement experiments.

5. CONCLUSIONS.

The results that were most relevant to the goals of this study, as well as some tentative conclusions, are presented below.

1. The memory aids decreased the number of ground-to-air transmissions, including both altitude and heading assignments.

2. The memory aids reduced the number of handoff misses, but not the number of conflict errors.

3. Controllers' and the expert observer's ratings of performance indicated that the memory aids slightly increased perceived performance in low traffic scenarios, but slightly decreased perceived performance in high traffic scenarios.

4. Controllers' ratings of SA were slightly higher with the memory aids, but the SAGAT measures indicated that the memory aids had no effect on SA.

5. Controllers' ratings of workload, as measured by ATWIT, were not affected by the memory aids, but were directly related to the taskload in scenarios.

6. High traffic scenarios were related to decreased performance, lowered SA, and increased workload relative to low traffic scenarios.

7. Controllers' ratings of SA did not correlate with the objective measures of SA obtained by the present SAGAT procedure.

8. Controllers' comments indicated that they did not have enough time to learn the memory aids completely.

9. Controllers' comments indicated that accuracy in the present SAGAT procedure required them to remember aircraft call signs which they felt was not necessary in actual ATC.

10. Although some disadvantages were mentioned, most controllers felt the memory aids were helpful and ratings of effectiveness, usability, and acceptability were favorable.

11. The memory aids did not improve controller performance, increase SA, or reduce workload as much as expected. However, the memory aids may have been more effective if controllers had more time to learn the experimental procedures.

12. The present study did not confirm the utility of the SAGAT procedure as an accurate objective measure of SA. However, the present experiment involved a unique format for the SAGAT questions and the results have provided direction for improvement of the technique in future research.

REFERENCES

American Psychological Association, Human Factors and Ergonomics Society, & Federation of Behavioral, Psychological, and Cognitive Sciences (1994). Written Testimony for the U. S. Senate Committee on Appropriations, Subcommittee on Transportation on the subject of the Fiscal Year 1995 Appropriations for the Federal Aviation Administration.

Baddeley, A. D. (1986). Working memory. Oxford: Clarendon Press.

Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized work lists. Journal of Verbal Learning and Verbal Behavior, 8, 323–343.

Buckley, E. P., DeBaryshe, B. D., Hitchner, N., & Kohn, P. (1983). Methods and measurements in real-time air traffic control system simulation (DOT/FAA/CT-83/26). Atlantic City, NJ: DOT/FAA Technical Center.

Endsley, M. R. (1987). SAGAT: A methodology for the measurement of situation awareness (NOR DOC 87-83). Hawthorne, CA: Northrop Corporation.

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

Endsley, M. R. (1989). A methodology for the objective measurement of situation awareness. In Situational awareness in aerospace operations (pp 1-9) Copenhagen, Denmark: NATO-AGARD.

Endsley, M. R. (1990). Predictive utility of and objective measure of situation awareness. Proceedings of the Human Factors Society 34th Annual Meeting. Santa Monica, CA: Human Factors Society.

Endsley, M. R., & Bolstad, C. A. (1993). Human capabilities and limitations in situation awareness. Paper presented at the NATO-AGARD symposium on Combat Automation for Airborne Weapons Systems: Man Machine Interface Trends and Technologies, Edinburgh, Scotland.

Federal Aviation Administration (1987). Profile of operational errors in the national airspace system, calendar year 1986. Washington, DC.

Federal Aviation Administration (1988). Profile of operational errors in the national airspace system, calendar year 1987. Washington, DC.

Federal Aviation Administration (1989). Air traffic control (DOT/FAA Handbook 7110.65F). Washington, DC: U.S. Government Printing Office.

Federal Aviation Administration (1994). Administrator's fact book (AMS-400). Washington, DC: U.S. Government Printing Office.

Finklman, J. M., & Kirchner, C. (1980). An information processing interpretation of air traffic control stress. Human Factors, 22, 561–567.

Flach, J. M. (1994). Situation awareness: The emperor's new clothes. Paper presented at the conference on Automation Technology, Washington, DC.

Garland, D. J., Stein, E. S., Blanchard, J. W., & Wise, J. A. (1992). Situational awareness in the future air traffic control environment. Proceedings of the 37th Air Traffic Control Association Conference Atlantic City, NJ.

Garland, D. J., Stein, E. S., Blanchard, J. W., & Wise, J. A. (1993). Situational awareness in air traffic control, a critical yet neglected phenomenon. In D. J. Garland and J. A. Wise (Eds.), Human factors and advanced aviation technologies. Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Gromelski, S., Davidson, L., & Stein, E. S. (1992). Controller memory enhancement: Field facility concepts and techniques (DOT/FAA/CT-TN92/7). Atlantic City, NJ: DOT/FAA Technical Center.

Hitchcock, L. (1993). Yours, mine, and ours: Some observations on the metaphysics of situation awareness. In R. Gibson, D. J. Garland and J. Koonce (Eds.), Situation Awareness in Complex Systems. Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Hopkin, V. D. (1991). Automated flight strip usage: Lessons from the functions of paper strips. Paper presented at the conference on Challenges in Aviation Human Factors: The National Plan, Washington, DC.

Jackson, A. (1989). The functionality of flight strips. Report to the U.K. Civil Aviation Authority: Royal Signals and Radar Establishment, Farnborough, U.K.

Kinney, G. C., Spahn, M. J., & Amato, R. A. (1977). The human element in air traffic control: Observations and analyses of the performance of controllers and supervisors in providing separation services (Report No. MTR-7655). McLean, VA: The MITRE Corporation.

Kirchner, J. H., & Laurig, W. (1971). The human operator in air traffic control. Ergonomics, 14, 549–556.

Klapp, S. T., Marshburn, E. A., & Lester, P. T. (1983). Short-term memory does not involve the "working memory" of information processing: The demise of a common assumption. Journal of Experimental Psychology: General, 112, 240–264.

Koriat, A., Ben-Zur, H., & Nussbaum, A. (1990). Encoding information for future action: Memory for to-be-performed tasks versus memory for to-be-recalled tasks. Memory and Cognition, 18, 568–578.

Means, B., Mumaw, R., Roth, C., Schlager, M., McWilliam, E., Gagne, E., Rice, V., Rosenthal, D., & Heon, S. (1988). ATC training analysis study: Design of the next generation ATC training system (OPM Work Order No. 342-036). Richmond, VA: HumRRO International.

Miller, G. (1956). The magic number seven plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81–97.

Moray, N. (1986). Monitoring behavior and supervisory control. In J. Rasmussen and W. B. Rouse (Eds.), Handbook of perception and human performance: Volume II, Cognitive processes and performance (pp. 40–51). New York: Wiley.

Norman, D. (1992). Turn signals are the facial expressions of automobiles. Reading, MA: Addison-Wesley.

Operational Error Analysis Work Group (1987). Actions to implement recommendations of April 17, 1987. Unpublished manuscript, Federal Aviation Administration, Washington, DC.

Rasmussen, J. (1987). Cognitive control and human error mechanisms. In J. Rasmussen, K. Duncan, and J. Leplat (Eds.), New technology and human error (pp. 53–60). New York John Wiley and Sons.

Rasmussen, J. (1988). Information technology: A challenge to the human factors society Human Factors Society Bulletin, 31, 1–3.

Rasmussen, J., & Lind, M. (1982). A model of human decision-making in complex systems and its use for design of system control strategies (Report No. RISO-M-2349). Roskilde, Denmark: RISO National Laboratory.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. The International Journal of Aviation Psychology, 1(1), 45–57.

Senders, J. W., & Moray, N. P. (1991). Human error: Cause, prediction, and reduction. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sperandio, J. M. (1971). Variations of operator's strategies in regulating the effects of workload. Ergonomics, 14, 571–577.

Spettel, C. M., & Liebert, R. M. (1986). Training for safety in automated person-machine systems. American Psychologist, 41, 545–550.

Stager, P., & Hameluck, D. (1990). Ergonomics and air traffic control. Ergonomics, 33(4), 493–499.

Stein, E. S. (1985). Air traffic controller workload: An examination of workload probe (DOT/FAA/CT-TN84/24). Atlantic City, NJ: DOT/FAA Technical Center.

Stein, E. S. (1991). Air traffic controller memory—a field survey (DOT/FAA/CT-TN90/60). Atlantic City, NJ: DOT/FAA Technical Center.

Stein, E. S., & Bailey, J. (1989). The controller memory handbook (DOT/FAA/CT-TN89/58). Atlantic City, NJ: DOT/FAA Technical Center.

Tenny, Y. J., Adams, M. J., Pew, R. W., Huggins, A. F., & Rogers, W. H. (1992). A principled approach to the measurement of situation awareness in commercial aviation (Contract No. 4451). Langley, VA: NASA Langley Research Center.

Thomas, D. D. (1985). ATC in transition, 1956–1963. Journal of Air Traffic Control, 30–38.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. Psychological Review, 69, 344–354.

UFA, Inc. (1992). ATCoach [Computer software]. Lexington, MA: UFA, Inc.

Vingelis, P. J., Schaeffer, E., Stringer, P., Gromelski, S., & Ahmed, B., (1990). Air traffic controller memory enhancement: Literature review and proposed memory aids (DOT/FAA/CT-TN90/38). Atlantic City, NJ: DOT/FAA Technical Center.

Vortac, O. U. (1991). Cognitive factors in the use of flight progress strips: Implications for automation. Unpublished technical report, University of Oklahoma, Norman, OK.

Vortac, O. U., Edwards, M. B., Fuller, D. K., & Manning, C. A. (1994). Automation and cognition in air traffic control: An empirical investigation (DOT/FAA/AM-94/3). Oklahoma City, OK: FAA Aeromedical Institute.

Vortac, O. U., Edwards, M. B., Jones, J. P. Manning, C. A., & Rotter, A. J. (1992). En route air traffic controllers use of flight progress strips: A graph theoretical analysis (DOT/FAA AM-92/31). Oklahoma City, OK: FAA Civil Aeromedical Institute.

Warm, J. S., & Dember, W. M. (1986). Awake at the switch. Psychology Today, 20, 46–53.

Zingale, C., Gromelski, S., & Stein, E. S. (1992). Preliminary studies of planning and flight strip use as air traffic controller memory aids (DOT/FAA/CT-TN92/22). Atlantic City, NJ: DOT/FAA Technical Center.

Zingale, C., Gromelski, S., Ahmed, B., & Stein, E. S. (1993). Influence of individual experience and flight strips on air traffic controller memory/situational awareness (DOT/FAA/CT-TN93/31). Atlantic City, NJ: DOT/FAA Technical Center.

APPENDIX A
QUESTIONNAIRES

INITIAL QUESTIONNAIRE

Participant Code Date

INSTRUCTIONS

The purpose of this questionnaire is to obtain information concerning your background and feelings about this study in order to better understand your performance during the course of the experiment. This information will be used to describe the participants in this study as a group and then relate the group's characteristics to how you perform and what you tell us during the experiment. So that your identity can remain anonymous, your actual name should not be written on this form. Instead, your data will be identified by a participant code known only to yourself and the experimenters.

1) What is your age, in years?

_____ years

2) How many years have you actively controlled traffic?

years

3) How many months in the past year have you actively controlled traffic?

_ months

4) What is your current position as an air traffic controller?

❑ Developmental ❑ Full Performance Level ❑ Other

5) Do you wear corrective lenses?

❑ Yes ❑No

If yes, do you have them with you?

❑ Yes ❑No

6) Circle the number which best describes your current skill as an air traffic controller.

1	2	3	4	5	6	7	8	9	10
Not Very Skilled									Extremely Skilled

7) Circle the number which best describes your motivation to participate in this study.

1	2	3	4	5	6	7	8	9	10
Not Very Motivated									Extremely Motivated

8) Circle the number which best describes your current state of health.

	2	3	4	5	6	7	8	9	10
Not Very Healthy									Extremely Healthy

9) Circle the number which best describes your level of video game experience.

1	2	3	4	5	6	7	8	9	10
Not Very Experienced									Extremely Experienced

POST-SCENARIO QUESTIONNAIRE

Participant Code ________ Date ________

Scenario Code

INSTRUCTIONS

The purpose of this questionnaire is to obtain information concerning different aspects of the scenario just completed. This information will be used to determine how the manipulations in this scenario affect your opinions. As you answer each question, feel free to use the entire numerical scale. Please be as honest and as accurate as you can. So that your identity can remain anonymous, your actual name should not be written on this form. Instead, your data will be identified by a participant code known only to yourself and the experimenters.

1) Please circle the number below which best describes how hard you were working during this scenario.

Very Low Workload (All tasks were accomplished easily and quickly)	1 2 3
Moderate Workload (The chances for error or omission were low)	
Relatively High Workload (The chances for some error or omission were relatively high)	
Very High Workload (It was not possible to accomplish all tasks properly)	

2) How well did you control traffic in this scenario?

1	2	3	4	5	6	7	8	9	10
Not Very Well									Extremely Well

3) Please rate your overall situational awareness during this scenario.

1	2	3	4	5	6	7	8	9	10
Not Very Aware									Extremely Aware

4) Please rate your situational awareness for current aircraft locations during this scenario.

1	2	3	4	5	6	7	8	9	10
Not Very Aware									Extremely Aware

5) Please rate your situational awareness for projected aircraft locations during this scenario.

1	2	3	4	5	6	7	8	9	10
Not Very Aware									Extremely Aware

6) Please rate your situational awareness for potential safety violations during this scenario.

1	2	3	4	5	6	7	8	9	10
Not Very Aware									Extremely Aware

7) How difficult was this scenario?

1	2	3	4	5	6	7	8	9	10
Not Very Difficult									Extremely Difficult

OBSERVER CHECKLIST

Participant Code ______________________ Date ________

Scenario Code

INSTRUCTIONS

The purpose of this questionnaire is to obtain information concerning the participant's style of controlling traffic in the scenario just completed. Controllers may use different techniques during the scenario, however, you should indicate the participant's most frequent action unless requested otherwise. Also, you should wait until the scenario has finished to indicate your final decisions, although please feel free to make preliminary notes during the course of the scenario.

Scanning Techniques

1) Which statement best describes the controller's scanning pattern?

❑Hot Spot Scanning

(Controller concentrates attention to a location on the scope that has the heaviest volume of traffic.)

❑Entire Scope Scanning

(Controller follows the radar sweep and takes action accordingly.)

❑Center of Scope to Outer Edges

(Controller starts at airport and works outward to scopes edge.)

❑Scanning Divided Between Scope and Strips

(Controller uses one of the techniques above (please indicate) but scanning pattern also includes using flight progress strips as memory aid.)

❑Other Scanning Techniques

____

____

2) What was the extent of reliance placed on displayed data?

❑ Less than 50% ❑ 50% to 75% ❑ More than 75%

3) Did the controller constantly scan the scope updating information as it became available?

❑ Yes ❑ No

4) Did the controller use flight progress strips when traffic conditions were moderate to heavy?

❑ Yes ❑ No

Decision Making Processes

1) Which statement best describes the controller's decision making process?

- ❑ Controller immediately took action to resolve the potential of an impending conflicts.
- ❑ Controller projected flight plan conflict, planned appropriate action to eliminate conflict and then acted accordingly.
- ❑ Controller projected flight plan conflict, waited until it was apparent that a conflict was imminent and then took corrective action.
- ❑ Controller either was not aware of an impending conflict or waited until the last minute to take corrective action.
- ❑ Controller was overwhelmed with traffic and took corrective action in an unorganized, unplanned manner. The actions taken were not always prioritized.
- ❑ Other Decision Making Processes

____

____

Communication Skills

The controller used incorrect phraseology.

❑ Rarely ❑ Occasionally ❑ Frequently

2) The controller used extraneous transmissions.

❑ Rarely ❑ Occasionally ❑ Frequently

3) The controller repeated control instructions.

❑ Rarely ❑ Occasionally ❑ Frequently

4) The controller used improper voice inflections.

❑ Rarely ❑ Occasionally ❑ Frequently

Separation Techniques

1) Under light traffic conditions, which of the following techniques were used to separate aircraft? If more than one technique was used, please indicate which ones.

❑ Radar vectors ❑ Altitude separation ❑ Speed control

2) Under moderate traffic conditions, which of the following techniques were used to separate aircraft? If more than one technique was used, please indicate which ones.

❑ Radar vectors ❑ Altitude separation ❑ Speed control

3) Under heavy traffic conditions, which of the following techniques were used to separate aircraft? If more than one technique was used, please indicate which ones.

❑ Radar vectors ❑ Altitude separation ❑ Speed control

Other Observed Techniques

1) Which of these techniques best describe the controller's overall approach to working traffic during the scenario? If more than one technique was used, please indicate which ones.

❑ Completed the Transaction

(Initially performed most tasks associated with controlling the aircraft.)

❑ Standardization

(Used the same route, speed and control instructions whenever possible.)

❑ Chunking

(Placed control instructions into familiar units and recalled them as one chunk.

❑ Preplanning

(Reviewed flight progress strips for pending traffic and formulated a tentative plan prior to actually working the traffic.)

❑ Visualization

(Envisioned the projected route of all traffic under their control and mentally imaged the flight's progress through their airspace.)

❑ Prioritization

(The ability to correctly determine which aircraft rate the highest priority when issuing corrective control instructions.)

❑ Timeliness

(The ability to issue control instructions in an effective (timely) manner.)

❑ Professionalism

(The manner in which the controller assumed control of the position. This area deals with voice inflection, speech rate and common courtesy for others sharing a common voice frequency.)

❑ Randomization

(Controller did not have a definite approach or game plan for working the traffic. This technique for controlling traffic became more obvious under moderate or heavy traffic conditions.)

❑ Additional Comments

____

____

Subjective Analysis of Performance

1) Please circle the number below which best describes how hard the controller was working during this scenario.

Very Low Workload	1
(All tasks were accomplished easily and quickly)	2
	3
Moderate Workload	
(The chances for error or omission were low)	
Relatively High Workload	
(The chances for some error or omission were relatively high)	
Very High Workload	
(It was not possible to accomplish all tasks properly)	
	12

2) How well did the participant control traffic during this scenario?

1	2	3	4	5	6	7	8	9	10
Not Very Well									Extremely Well

3) How often did the participant use the arrival and departure procedures during this scenario?

1	2	3	4	5	6	7	8	9	10
Not Very Often									Extremely Often

4) How often did the participant use "complete the transaction" communications during this scenario?

1	2	3	4	5	6	7	8	9	10
Not Very Often									Extremely Often

5) How well did the controller utilize time during this scenario?

1	2	3	4	5	6	7	8	9	10
Not Very Well									Extremely Well

6) How well was the controller able to recognize his/her traffic capacity limitations during this scenario?

1	2	3	4	5	6	7	8	9	10
Not Very Well									Extremely Well

INFORMATION IMPORTANCE FORM

Participant Code Date

INSTRUCTIONS

The purpose of this questionnaire is to obtain responses concerning the importance of different kinds of aircraft information. This information will be used to relate your importance ratings to your situational awareness. As you answer each question, feel free to use the entire numerical scale. Please be as honest and as accurate as you can. So that your identity can remain anonymous, your actual name should not be written on this form. Instead, your data will be identified by a participant code known only to yourself and the experimenters.

Please rate the importance of the following aircraft information for controller performance.

1 - Not Very Important 10 - Extremely Important

Aircraft Call Sign	1	2	3	4	5	6	7	8	9	10
Aircraft Type	1	2	3	4	5	6	7	8	9	10
Aircraft Beacon Code	1	2	3	4	5	6	7	8	9	10
Controller Ownership	1	2	3	4	5	6	7	8	9	10
Entry Altitude	1	2	3	4	5	6	7	8	9	10
Entry Airspeed	1	2	3	4	5	6	7	8	9	10
Entry Fix	1	2	3	4	5	6	7	8	9	10
Exit Altitude	1	2	3	4	5	6	7	8	9	10
Exit Airspeed	1	2	3	4	5	6	7	8	9	10
Exit Fix	1	2	3	4	5	6	7	8	9	10
Arrival Airport (within Sector)	1	2	3	4	5	6	7	8	9	10
Departure Airport (within Sector)	1	2	3	4	5	6	7	8	9	10
Current Altitude		2	3	4	5	6	7	8	9	10
Current Airspeed	1	2	3	4	5	6	7	8	9	10
Current Heading	1	2	3	4	5	6	7	8	9	10
Current Aircraft Location	1	2	3	4	5	6	7	8	9	10
Most Recently Assigned Altitude	1	2	3	4	5	6	7	8	9	10
Most Recently Assigned Airspeed	1	2	3	4	5	6	7	8	9	10
Most Recently Assigned Heading	1	2	3	4	5	6	7	8	9	10
Aircraft Holding/Spinning	1	2	3	4	5	6	7	8	9	10
Aircraft Waiting for Handoff/Release	1	2	3	4	5	6	7	8	9	10
Aircraft Near Exit Fix/Arrival Airport	1	2	3	4	5	6	7	8	9	10
Density of Aircraft on Radar Display	1	2	3	4	5	6	7	8	9	10

FINAL QUESTIONNAIRE

Participant Code Date

INSTRUCTIONS

The purpose of this questionnaire is to obtain feedback from you concerning different aspects of the experiment. This information will be used to improve our simulation in the future. In addition to ratings, you will be asked to make comments on some of the questions. Even if your ratings are rather favorable, you may wish to make further comments. If you feel you have any helpful ideas regarding this experiment, we would like to hear from you! So that your identity can remain anonymous, your actual name should not be written on this form. Instead, your data will be identified by a participant code known only to yourself and the experimenters.

1) In general, how realistic was the simulation?

1	2	3	4	5	6	7	8	9	10
Not Very Realistic									Extremely Realistic

If appropriate, please comment on why the simulation was not realistic?

____

____

____

____

2) How accurately did the simulation depict Atlantic City airspace?

1	2	3	4	5	6	7	8	9	10
Not Very Accurate									Extremely Accurate

If appropriate, please comment on why Atlantic City airspace was not accurate?

____

____

____

____

3) How accurately did low taskload scenarios simulate Atlantic City traffic during a moderately busy day?

1	2	3	4	5	6	7	8	9	10
Not Very Accurate									Extremely Accurate

If appropriate, please comment on why low taskload scenarios were not accurate?

____

____

____

____

4) How accurately did high taskload scenarios simulate Atlantic City traffic during an extremely busy day?

1	2	3	4	5	6	7	8	9	10
Not Very Accurate									Extremely Accurate

If appropriate, please comment on why high taskload scenarios were not accurate?

____

____

____

____

5) How useful were the experimental arrival procedures for improving your performance?

1	2	3	4	5	6	7	8	9	10
Not Very Useful									Extremely Useful

If appropriate, please try to describe why the experimental arrival procedures were not useful for improving your performance?

____

____

____

____

6) How useful were the experimental departure procedures for improving your performance?

1	2	3	4	5	6	7	8	9	10
Not Very Useful									Extremely Useful

If appropriate, please try to describe why the experimental departure procedures were not useful for improving your performance?

____

____

____

____

7) How useful was the "completing the transaction" technique for improving your performance?

1	2	3	4	5	6	7	8	9	10
Not Very Useful									Extremely Useful

If appropriate, please try to describe why the "completing the transaction" technique was not useful for improving your performance?

____

____

____

____

8) How useful were the experimental arrival procedures for increasing your situational awareness?

1	2	3	4	5	6	7	8	9	10
Not Very Useful									Extremely Useful

If appropriate, please try to describe why the experimental arrival procedures were not useful for improving your situational awareness?

____

____

____

____

9) How useful were the experimental departure procedures for increasing your situational awareness?

1	2	3	4	5	6	7	8	9	10
Not Very Useful									Extremely Useful

If appropriate, please try to describe why the experimental departure procedures were not useful for improving your situational awareness?

____

____

____

____

10) How useful was the “completing the transaction” technique for increasing your situational awareness?

1	2	3	4	5	6	7	8	9	10
Not Very Useful									Extremely Useful

If appropriate, please try to describe why the “completing the transaction” technique was not useful for improving your situational awareness?

____

____

____

____

11) How useful were the experimental arrival procedures for reducing your air traffic workload?

1	2	3	4	5	6	7	8	9	10
Not Very Useful									Extremely Useful

If appropriate, please try to describe why the experimental arrival procedures were not useful for reducing your air traffic workload?

____

____

____

____

12) How useful were the experimental departure procedures for reducing your air traffic workload?

1	2	3	4	5	6	7	8	9	10
Not Very Useful									Extremely Useful

If appropriate, please try to describe why the experimental departure procedures were not useful for reducing your air traffic workload?

____

____

____

____

13) How useful was the "completing the transaction" technique for reducing your air traffic workload?

1	2	3	4	5	6	7	8	9	10
Not Very Useful									Extremely Useful

If appropriate, please try to describe why the "completing the transaction" technique was not useful for reducing your air traffic workload?

____

____

____

____

14) Please rate how easy the experimental arrival procedures were to use.

1	2	3	4	5	6	7	8	9	10
Not Very Easy									Extremely Easy

If appropriate, please try to describe why the experimental arrival procedures were not easy to use?

____

____

____

____

15) Please rate how easy the experimental departure procedures were to use.

1	2	3	4	5	6	7	8	9	10
Not Very Easy									Extremely Easy

If appropriate, please try to describe why the experimental departure procedures were not easy to use?

____

____

____

____

16) Please rate how easy the "completing the transaction" technique was to use.

1	2	3	4	5	6	7	8	9	10
Not Very Easy									Extremely Easy

If appropriate, please try to describe why the "completing the transaction" technique was not easy to use?

____

____

____

____

17) Please rate your willingness to use the experimental arrival procedures.

1	2	3	4	5	6	7	8	9	10
Not Very Willing									Extremely Willing

If appropriate, please comment on why you were not willing to use the experimental arrival procedures?

____

____

____

____

18) Please rate your willingness to use the experimental departure procedures.

1	2	3	4	5	6	7	8	9	10
Not Very Willing									Extremely Willing

If appropriate, please comment on why you were not willing to use the experimental departure procedures?

____

____

____

____

19) Please rate your willingness to use the "completing the transaction" technique.

1	2	3	4	5	6	7	8	9	10
Not Very Willing									Extremely Willing

If appropriate, please comment on why you were not willing to use the "completing the transaction" technique?

____

____

____

____

20) To what extent did the ATWIT probe technique interfere with your performance?

1	2	3	4	5	6	7	8	9	10
Not Very Much									A Great Deal

21) To what extent did the SAGAT probe technique interfere with your performance?

1	2	3	4	5	6	7	8	9	10
Not Very Much									A Great Deal

22) If you have any ideas regarding memory aids for air traffic controllers, please describe your ideas below.

23) If you have any other comments or suggestions regarding this experiment, please write your ideas below.

APPENDIX B
SAGAT MATERIALS

SITUATIONAL AWARENESS QUESTIONNAIRE

Participant Code Date

Scenario Code

INSTRUCTIONS

Most controller's think of situational awareness as the ability to maintain a "mental picture" of the air traffic depicted on the radar display. Although this definition captures the essence of situational awareness, a more complete description involves an awareness of these three aspects of the air traffic environment: current locations of aircraft, projected locations of aircraft and an understanding of how this information relates to air traffic control system goals. The following questions are intended to assess both critical and less important information regarding these three aspects of your situational awareness. Try to proceed through the questions as quickly as possible making your best guess if you are not certain of the correct answer. So that your identity can remain anonymous, your actual name should not be written on this form. Instead, your data will be identified by a participant code known only to yourself and the experimenters.

(Always Presented - Preliminary Question)

1) Has the specified aircraft been cleared for a runway approach?

❑ No ❑ Yes

(Always Presented - Preliminary Question)

2) Has the specified aircraft been handed off to another controller?

❑ No ❑ Yes

(Presented for All Aircraft Flight Plans)

3) How many aircraft (controlled by you) are within a 5 mile radius and below the specified aircraft?

_____ Aircraft

(Presented for All Aircraft Flight Plans)

4) How many aircraft (controlled by you) are within a 5 mile radius and above the specified aircraft?

_____ Aircraft

(Presented for All Aircraft Flight Plans)

5) How many aircraft (controlled by you) are within a 10 mile radius and at the same altitude as the specified aircraft?

_____ Aircraft

(Presented for All Aircraft Flight Plans)

6) What is the current altitude (in feet) of the specified aircraft?

Current Altitude _____ feet

(Presented for All Aircraft Flight Plans)

7) What is the current airspeed (in knots) of the specified aircraft?

Current Airspeed _____ knots

(Presented for All Aircraft Flight Plans)

8) What is the current heading or fix of the specified aircraft?

Current Heading __________

(Presented for All Aircraft Flight Plans)

9) What was the most recently assigned altitude (in feet) for the specified aircraft?

❑No Assignment Made Assigned Altitude _____ feet

(Presented for All Aircraft Flight Plans)

10) What was the most recently assigned airspeed (in knots) for the specified aircraft?

❑No Assignment Made Assigned Airspeed _____ knots

(Presented for All Aircraft Flight Plans)

11) What was the most recently assigned heading or fix for the specified aircraft according?

❑ No Assignment Made Assigned Heading __________

(Presented for Arrival Aircraft Only)

12) What will be the arrival airport of the specified aircraft?

Arrival Airport __________

(Presented for Arrival and Overflight Aircraft Only)

13) What will be the entry fix of the specified aircraft?

Entry Fix __________

(Presented for Departure Aircraft Only)

14) What will be the departure airport of the specified aircraft?

Departure Airport __________

(Presented for Departure and Overflight Aircraft Only)

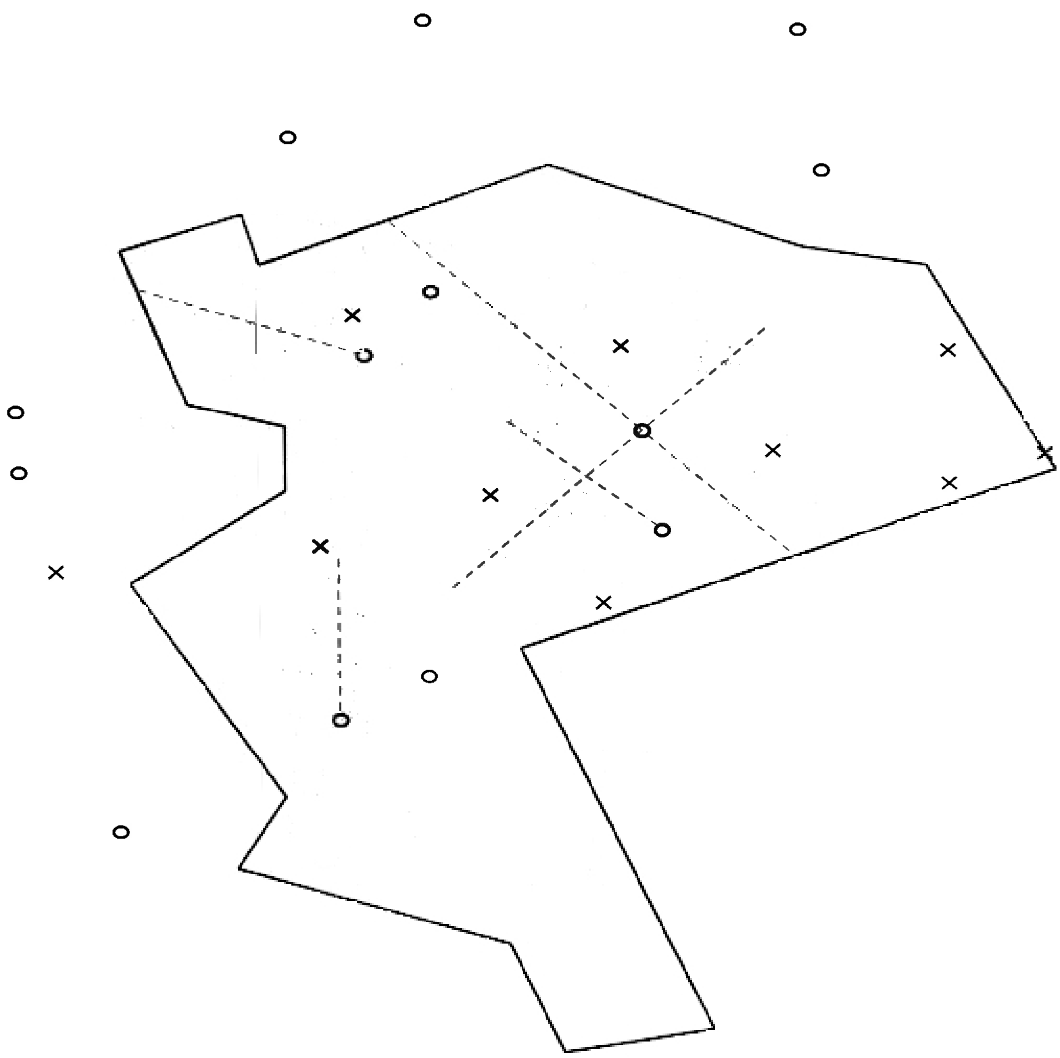
15) What will be the exit fix of the specified aircraft?

Exit Fix __________

SPATIAL LOCATION MAP

INSTRUCTIONS
Listed are eight aircraft call signs and a number associated with each aircraft. You should place the number that represents each aircraft in the proper location on the map provided below.

1 - N272A
2 - N72XG
3 - N2814R
4 - PDT3133
5 - N41JA
6 - N40ZG
7 - RDB743
8 - 2147Z



APPENDIX C
SUMMARY TABLES

TABLE C1. MEANS AND STANDARD DEVIATIONS FOR THE NUMBER OF CONFLICTS (NCNF)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	0.06 (0.17)	0.00 (0.00)
	0.31 (0.48)	0.50 (0.58)

TABLE C2. MEANS AND STANDARD DEVIATIONS FOR THE NUMBER OF HANDOFF MISSES (NHOMISS)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	0.13 (0.29)	0.03 (0.13)
	0.06 (0.17)	0.03 (0.13)

TABLE C3. MEANS AND STANDARD DEVIATIONS FOR THE NUMBER OF ALTITUDE ASSIGNMENTS (NALT)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	13.25 (1.91)	5.03 (1.19)
	23.34 (2.54)	12.13 (2.30)

TABLE C4. MEANS AND STANDARD DEVIATIONS FOR THE NUMBER OF HEADING ASSIGNMENTS (NHDG)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	15.59 (2.54)	3.47 (0.76)
	28.38 (3.07)	11.66 (2.94)

TABLE C5. MEANS AND STANDARD DEVIATIONS FOR THE NUMBER OF SPEED ASSIGNMENTS (NSPD)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	0.03 (0.13)	0.25 (0.45)
	0.38 (0.72)	0.34 (0.51)

TABLE C6. MEANS AND STANDARD DEVIATIONS FOR THE NUMBER OF GROUND-TO-AIR TRANSMISSIONS (NPTT)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	36.81 (7.95)	28.37 (3.74)
	60.69 (8.75)	50.23 (6.26)

TABLE C7. MEANS AND STANDARD DEVIATIONS FOR THE CUMULATIVE AVERAGE OF SYSTEM ACTIVITY (CMAV)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	1.49 (0.29)	1.59 (0.34)
	5.21 (0.42)	5.30 (0.40)

TABLE C8. MEANS AND STANDARD DEVIATIONS FOR THE NUMBER OF FLIGHTS COMPLETED (NCOMP)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	9.50 (0.88)	9.59 (0.78)
	14.09 (0.99)	14.59 (1.94)

TABLE C9. MEANS AND STANDARD DEVIATIONS FOR THE PERCENTAGE OF POINTS OBTAINED IN THE INFORMATION RECALL PHASE (SGTIR)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	72.93 (10.40)	73.82 (14.91)
	61.06 (17.09)	60.73 (18.79)

TABLE C10. MEANS AND STANDARD DEVIATIONS FOR THE PERCENTAGE OF POINTS OBTAINED IN THE SPATIAL LOCATION PHASE (SGTSL)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	73.96 (20.56)	69.24 (22.62)
	64.04 (17.17)	55.35 (17.81)

TABLE C11. MEANS AND STANDARD DEVIATIONS FOR THE AIR TRAFFIC WORKLOAD INPUT TECHNIQUE (ATWIT)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	2.31 (0.86)	2.38 (0.75)
	4.62 (1.08)	4.59 (1.15)

TABLE C12. MEANS AND STANDARD DEVIATIONS FOR THE CONTROLLER'S WORKLOAD RATING (PSQWRK)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	3.34 (1.19)	3.22 (1.06)
	7.94 (1.60)	7.63 (1.54)

TABLE C13. MEANS AND STANDARD DEVIATIONS FOR THE CONTROLLER'S PERFORMANCE RATING (PSQPFM)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	8.38 (1.13)	8.59 (1.16)
	7.16 (1.11)	6.78 (1.45)

TABLE C14. MEANS AND STANDARD DEVIATIONS FOR THE CONTROLLER'S OVERALL SITUATIONAL AWARENESS RATING (PSQSAW)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	8.31 (1.17)	8.81 (1.00)
	6.91 (1.40)	7.16 (1.51)

TABLE C15. MEANS AND STANDARD DEVIATIONS FOR THE CONTROLLER'S CURRENT AIRCRAFT LOCATION AWARENESS RATING (PSQCAL)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	7.75 (1.49)	7.81 (1.38)
	6.72 (1.72)	6.53 (1.54)

TABLE C16. MEANS AND STANDARD DEVIATIONS FOR THE CONTROLLER'S PROJECTED AIRCRAFT LOCATION AWARENESS RATING (PSQPAL)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	8.16 (1.26)	8.16 (1.43)
	6.78 (1.78)	6.69 (2.06)

TABLE C17. MEANS AND STANDARD DEVIATIONS FOR THE CONTROLLER'S POTENTIAL SAFETY VIOLATIONS AWARENESS RATING (PSQPSV)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	8.94 (0.98)	9.19 (0.60)
	8.16 (1.14)	7.63 (1.95)

TABLE C18. MEANS AND STANDARD DEVIATIONS FOR THE CONTROLLER'S SCENARIO DIFFICULTY RATING (PSQDIF)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	3.13 (1.34)	2.88 (1.28)
	7.56 (1.20)	7.50 (1.37)

TABLE C19. MEANS AND STANDARD DEVIATIONS FOR THE EXPERT OBSERVER'S WORKLOAD RATING (OBSWRK)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	3.44 (1.01)	3.25 (0.52)
	8.53 (1.32)	8.31 (1.62)

TABLE C20. MEANS AND STANDARD DEVIATIONS FOR THE EXPERT OBSERVER'S PERFORMANCE RATING (OBSPFM)

TRAFFIC	NO MEMORY AIDS Mean (SD)	MEMORY AIDS Mean (SD)
	8.66 (0.91)	9.16 (0.70)
	7.17 (1.08)	6.84 (1.55)

APPENDIX D
ANOVA TABLES

TABLE D1. ANOVA RESULTS FOR THE NUMBER OF CONFLICTS (NCNF)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		0.039583		
MEMORY		0.010417	0.71	0.41
MEMORY*SUBJECT		0.014583		
TRAFFIC		0.375000	12.17	0.00
TRAFFIC*SUBJECT		0.029167		
MEMORY*TRAFFIC		0.041667	2.73	0.11
MEMORY*TRAFFIC*SUBJECT		0.015278		

TABLE D2. ANOVA RESULTS FOR THE NUMBER OF HANDOFF MISSES (NHOMISS)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		0.100000		
MEMORY		0.062500	5.00	0.04
MEMORY*SUBJECT		0.012500		
TRAFFIC		0.015625	1.00	0.33
TRAFFIC*SUBJECT		0.015625		
MEMORY*TRAFFIC		0.015625	1.00	0.33
MEMORY*TRAFFIC*SUBJECT		0.015625		

TABLE D3. ANOVA RESULTS FOR THE NUMBER OF ALITITUDE ASSIGNMENTS (NALT)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		7.225000		
MEMORY		1511.265625	369.45	0.00
MEMORY*SUBJECT		4.090625		
TRAFFIC		1181.640625	457.59	0.00
TRAFFIC*SUBJECT		2.582292		
MEMORY*TRAFFIC		36.000000	12.31	0.00
MEMORY*TRAFFIC*SUBJECT		2.925000		

TABLE D4. ANOVA RESULTS FOR THE NUMBER OF HEADING ASSIGNMENTS (NHDG)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		11.801822		
MEMORY		3327.847656	468.87	0.00
MEMORY*SUBJECT		7.097656		
TRAFFIC		1758.753906	375.22	
TRAFFIC*SUBJECT		4.687240		
MEMORY*TRAFFIC		84.410156	55.89	
MEMORY*TRAFFIC*SUBJECT		1.510156		

TABLE D5. ANOVA RESULTS FOR THE NUMBER OF SPEED ASSIGNMENTS (NSPD)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		0.341667		
MEMORY		0.140625	0.77	
MEMORY*SUBJECT		0.182292		
TRAFFIC		0.765625	3.18	
TRAFFIC*SUBJECT		0.240625		
MEMORY*TRAFFIC		0.250000	1.11	
MEMORY*TRAFFIC*SUBJECT		0.225000		

TABLE D6. ANOVA RESULTS FOR THE NUMBER OF GROUND-TO-AIR TRANSMISSIONS (NPTT)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		108.463333		
MEMORY		1396.837500	26.36	0.00
MEMORY*SUBJECT		52.989286		
TRAFFIC		8010.602781	1131.89	0.00
TRAFFIC*SUBJECT		7.077222		
MEMORY*TRAFFIC		17.604167	0.67	
MEMORY*TRAFFIC*SUBJECT		26.184524		

TABLE D7. ANOVA RESULTS FOR THE CUMULATIVE AVERAGE OF SYSTEM ACTIVITY (CMAV)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		0.160232		
MEMORY		0.130502	1.14	0.30
MEMORY*SUBJECT		0.114822		
TRAFFIC		221.451602	2220.11	0.00
TRAFFIC*SUBJECT		0.099748		
MEMORY*TRAFFIC		0.001139	0.01	0.94
MEMORY*TRAFFIC*SUBJECT		0.166419		

TABLE D8. ANOVA RESULTS FOR THE NUMBER OF FLIGHTS COMPLETED (NCOMP)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		1.991406		
MEMORY		1.410156	.17	0.30
MEMORY*SUBJECT		1.210156		
TRAFFIC		368.160156	603.39	0.00
TRAFFIC*SUBJECT		0.610156		
MEMORY*TRAFFIC		0.660156	0.29	0.60
MEMORY*TRAFFIC*SUBJECT		0.310156		

TABLE D9. ANOVA RESULTS FOR THE PERCENTAGE OF POINTS OBTAINED IN THE INFORMATION RECALL PHASE (SGTIR)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		444.431465		
MEMORY		1.262814	0.01	0.92
MEMORY*SUBJECT		120.310254		
TRAFFIC		2493.129727	10.99	0.00
TRAFFIC*SUBJECT		226.915793		
MEMORY*TRAFFIC		5.862452	0.03	0.86
MEMORY*TRAFFIC*SUBJECT		183.902118		

TABLE D10. ANOVA RESULTS FOR THE PERCENTAGE OF POINTS OBTAINED IN THE SPATIAL LOCATION PHASE (SGTSL)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		970.711585		
MEMORY		719.647689	2.65	0.12
MEMORY*SUBJECT		272.036856		
TRAFFIC		2266.117014	19.95	0.00
TRAFFIC*SUBJECT		113.595001		
MEMORY*TRAFFIC		62.984064	0.33	0.57
MEMORY*TRAFFIC*SUBJECT		189.853024		

TABLE D11. ANOVA RESULTS FOR THE AIR TRAFFIC WORKLOAD INPUT TECHNIQUE (ATWIT)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		2.253256		
MEMORY		0.009025	0.01	0.92
MEMORY*SUBJECT		0.812985		
TRAFFIC		82.128906	147.52	0.00
TRAFFIC*SUBJECT		0.556713		
MEMORY*TRAFFIC		0.032400	0.18	0.68
MEMORY*TRAFFIC*SUBJECT		0.183060		

TABLE D12. ANOVA RESULTS FOR THE CONTROLLER'S WORKLOAD RATING (PSQWRK)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT	15	3.795833		
MEMORY	1	0.765625	1.02	0.33
MEMORY*SUBJECT	15	0.748958		
TRAFFIC	1	324.000000	137.87	0.00
TRAFFIC*SUBJECT	15	2.350000		
MEMORY*TRAFFIC	1	0.140625	0.23	0.64
MEMORY*TRAFFIC*SUBJECT	15	0.607292		

TABLE D13. ANOVA RESULTS FOR THE CONTROLLER'S PERFORMANCE RATING (PSQPFM)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		3.526823		
MEMORY		0.097656	0.13	0.72
MEMORY*SUBJECT		0.755990		
TRAFFIC		36.753906	28.01	0.00
TRAFFIC*SUBJECT		1.312240		
MEMORY*TRAFFIC		1.410156	4.01	0.06
MEMORY*TRAFFIC*SUBJECT		0.351823		

TABLE D14. ANOVA RESULTS FOR THE CONTROLLER'S OVERALL SITUATIONAL AWARENESS RATING (PSQSAW)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		3.857292		
MEMORY		2.250000	3.46	0.08
MEMORY*SUBJECT		0.650000		
TRAFFIC		37.515625	31.29	0.00
TRAFFIC*SUBJECT		1.198958		
MEMORY*TRAFFIC		0.250000	0.27	0.61
MEMORY*TRAFFIC*SUBJECT		0.916667		

TABLE D15 ANOVA RESULTS FOR THE CONTROLLER'S CURRENT AIRCRAFT LOCATION AWARENESS RATING (PSQCAL)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		5.782292		
MEMORY		0.062500	0.03	0.87
MEMORY*SUBJECT		2.320833		
TRAFFIC		21.390625	28.56	0.00
TRAFFIC*SUBJECT	15	0.748958		
MEMORY*TRAFFIC	1	0.250000	0.40	0.54
MEMORY*TRAFFIC*SUBJECT	15	0.625000		

TABLE D16. ANOVA RESULTS FOR THE CONTROLLER'S PROJECTED AIRCRAFT LOCATION AWARENESS RATING (PSQPAL)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		5.983073		
MEMORY		0.035156	0.02	0.89
MEMORY*SUBJECT		1.876823		
TRAFFIC		32.347656	14.13	
TRAFFIC*SUBJECT	15	2.289323		
MEMORY*TRAFFIC	1	0.035156	0.04	0.85
MEMORY*TRAFFIC*SUBJECT	15	0.926823		

TABLE D17. ANOVA RESULTS FOR THE CONTROLLER'S POTENTIAL SAFETY VIOLATIONS AWARENESS RATING (PSQPSV)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		3.726823		
MEMORY		0.316406	0.31	
MEMORY*SUBJECT		1.024740		
TRAFFIC		21.972656	19.43	
TRAFFIC*SUBJECT		1.130990		
MEMORY*TRAFFIC		2.441406	4.44	
MEMORY*TRAFFIC*SUBJECT		0.549740		

TABLE D18. ANOVA RESULTS FOR THE CONTROLLER'S SCENARIO DIFFICULTY RATING (PSQDIF)

Source of Variation	DF	Mean Square	F-Ratio	Pr > F
SUBJECT		3.157292		
MEMORY		0.390625	0.39	
MEMORY*SUBJECT		0.998958		
TRAFFIC		328.515625	145.54	
TRAFFIC*SUBJECT		2.257292		
MEMORY*TRAFFIC		0.140625	0.45	
MEMORY*TRAFFIC*SUBJECT		0.315625		