DOT/FAA/AM-98/16

Office of Aviation Medicine
Washington, D.C. 20591

Aircraft Importance and Its Relevance to Situation Awareness

Scott D. Gronlund
Daryl D. Ohrt
Michael R. I'. Dougherty
Jennifer L. Perry
University of Oklahoma
Norman, OK 73019

Carol A. Manning
Civil Aeromedical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

May 1998

Final Report

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

U.S. Department of Transportation

Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-98/16		

4. Title and Subtitle	5. Report Date
Aircraft Importance and Its Relevance to Situation Awareness	May 1998
	6. Performing Organization Code
7. Author(s) Gronlund, S.D., Ohrt, D.D., Dougherty, M.R.P., perry, J.L. [1]; and Manning, C.A. [2]	8. Performing Organization Report No.
[1]University of Oklahoma Department of Psychology Norman, OK 73019 [2]FAA Civil Aeromedical Institute P.O. Box 25082 Oklahoma City, OK 73125	10. Work Unit No. (TRAIS)
9. Performing Organization Name and Address FAA Civil Aeromedical Institute P.O. Box 25082, Oklahoma City, OK 73125	11. Contract or Grant No. 13. Type of Report and Period Covered
12. Sponsoring Agency name and Address Office of Aviation Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	14. Sponsoring Agency Code

15. Supplemental Notes

This research was supported by Contract #DTFA-02-93-D-93088.

16. Abstract

We tested en route air traffic controllers (currently serving as instructors at the FAA Academy) to determine what they remember about the aircraft in their sector. We focused on memory for flight data (especially aircraft altitude and ground speed) and the position of the aircraft on the radar screen. Aircraft importance affected memory for flight data but not the highly accurate recall of the radar position of the aircraft. We hypothesize that controllers use their excellent memory for aircraft position to classify aircraft as important (potential traffic) or not, and better remember flight data about important aircraft (in particular, their exact altitude). The results have implications for improving techniques to assess situation awareness and interfaces to support it.

17. Key Words	18. Distribution Statement
Air Traffic Control, Situation Awareness, Memory	Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	14	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGMENTS

We are grateful to our subject-matter expert, Ken Shaver, without whom this experiment would have been impossible. We are also grateful to Dick Pollock for his cooperation. This research was supported by Contract #DTFA-02-93-D-93088 to the first author.

AIRCRAFT IMPORTANCE AND ITS RELEVANCE TO SITUATION AWARENESS

There is currently no agreed-upon methodology for assessing situation awareness (SA). Endsley (1995b) critically reviewed various methods, including physiological techniques, performance measures, and subjective techniques. The most commonly used method, according to Adams, Tenney, and Pew (1995), is the query technique (e.g., Endsley, 1987; Marshak, Kuperman, Ramsay, & Wilson, 1987). In this technique, the task simulation is suspended, the system displays are blanked, and the participant answers a series of questions about the situation. Query techniques tap the participant's ability to recall information about the situation from memory. According to Endsley (1995b, p 72), "SA, composed of highly relevant, attended to, and processed information, should be most receptive to recall." Endsley believes that the vast majority of a participant's SA can be assessed in this manner. To further explore the relationship between SA and memory, it is requisite to understand more about the role of memory in air traffic control.

The conclusions we can draw about the role of memory in air traffic control will be descriptive, rather than prescriptive. In other words, we can indicate what controllers remembered about the aircraft they were controlling, but we were unable to indicate what they *should* have remembered without an adequate measure of SA and/or performance. Current measures of air traffic control performance were too gross in their assessment to make this determination, summarizing performance over too wide a time-frame (e.g., Buckley, DeBaryshe, Hitchner, & Kohn, 1983-aircraft fuel consumption, delays; Vortac, Edwards, Fuller, & Manning, 1993-number of remaining control actions needed to exit aircraft from the sector). We need to correlate dynamic performance fluctuations with moment-to-moment changes in memory to be able to draw prescriptive conclusions. Nevertheless, to the extent that our participants maintained SA during our experiments, what they *do* remember might signal, in part, what they *should* remember.

If we knew what variables affected the recall of aircraft position on the radar and memory for flight data about an aircraft, we could make modifications to the query method. For example, currently, all aircraft are considered equivalent. We believe that there are some aircraft about which the controller should remember more, and other aircraft for which it would be acceptable (i.e., perfectly safe) that little was remembered. A second, more general goal, was to determine how memory was used by the air traffic controller; particularly, how aircraft position, altitude, and ground speed were represented. We begin by reviewing a study that first addressed many of these issues.

Means and associates (1988) conducted a study with three expert air traffic controllers. After controlling traffic for a period of time, the controllers completed a traffic drawing task in which they indicated the location of each aircraft on a paper copy of the sector map (see also Vortac and associates, 1993). Controllers performed exceedingly well on this task, correctly recalling upwards of 90% of the aircraft and correctly placing about 95% within 10 nautical miles of their actual positions. The ability to position the aircraft on the sector map stood in marked contrast to the recollection of many details regarding the aircraft. Means and associates (1988) found that controllers, when cued with the call sign, recalled only 28% of the aircraft types and only 6% of the ground speeds.

Means and associates (1988) proposed that the probability of recalling information about an aircraft was related to the *amount* of control exercised on the aircraft. This was operationalized as the number of control actions directed to a particular aircraft. There is ample support in the memory literature for the positive effect of frequency and repetition on memory (see Anderson, 1995). Means and associates (1988) found that twice as much flight data were recalled about "hot" aircraft (defined as aircraft for which controllers "exercised a great deal of control") than "cold" aircraft.

Gronlund, Dougherty, Ohrt, Thomson, Bleckley, Bain, Arnell, and Manning (1997) operationalized "hot" aircraft in two different ways: 1) by the number of control actions taken on an aircraft (as had Means and associates, 1988), and 2) by the number of interactions with an aircraft. An interaction was defined as any communication with an aircraft that did not result in a change to the aircraft's flight data; control actions were defined as any interaction that resulted in a change to the aircraft's altitude, speed, or heading. Gronlund and associates (1997) found no support for the hypothesis that the number of control actions (or the number of communications) affected the likelihood of recalling altitude information. However, because the assignment of aircraft to condition was determined by the actions of the controller in the Means and associates (1988) study, rather than being manipulated by the experimenter, as in the Gronlund experiment, some other factor might have been responsible for making an aircraft "hot" or "cold" in the Means study.

In a second experiment, Gronlund and associates (1997) tested whether that other factor was the *importance* of the aircraft in the scenario. We operationalized importance as the temporal proximity to the loss of separation. Thus, Means and associates (1988) might not have found that more was remembered about "hot" aircraft; instead, more might have been remembered about important aircraft (those potentially in conflict with another aircraft), which probably were the ones that received/required more control actions.

We proposed a hypothesis in which aircraft varied in importance in a graded fashion. However, the results of Gronlund and associates' (1997) second experiment indicated that this hypothesis needed refining. The manipulation of aircraft importance did affect memory for some aspects of the flight data of the aircraft (in particular, ground speed and relationship to sector, i.e., arrival, departure, or overflight), but not in the graded fashion that they expected. Furthermore, other aspects of the flight data, notably altitude and direction of flight, were not affected by aircraft importance. Interestingly, altitude and direction of flight are similar to two-dimensional position of the aircraft, in that they all relate to an aircraft's location in space (in this case, dynamic three-dimensional space).

Perhaps there are two distinct types of information underlying controller performance. 1) Memory for some flight data (e.g., speed) was affected by importance. However, rather than importance being graded, there may be only two categories: Important aircraft are those that are or might conflict, and unimportant aircraft will not conflict with another aircraft in the foreseeable future. 2) Information that pinpoints an aircraft's position in space (two-dimensional position on the Planned View Display or PVD, altitude, and direction of flight) was not affected by importance (nor by the "hot/cold" manipulation in Gronlund and associates' first experiment). We propose that this was because knowing the location of all aircraft was important to be able to make a classification of aircraft importance. These hypotheses were tested in the following experiment.

We sought to refine the importance hypothesis in two ways. First, we tested whether aircraft importance was binary-valued, rather than graded. If so, we would expect memory resources to be focused on the important aircraft, resulting in superior memory for their flight data (in particular, we will focus on altitude and ground speed). A second way we sought to refine the importance hypothesis was through the use of a better control condition. In Experiment 2 of Gronlund and associates (1997), the Not-traffic aircraft sometimes had been traffic for other aircraft earlier in the scenario. We suspect that many of the controllers would have classified some of these aircraft as "important" for that reason. This may explain why the flight data from the Not-traffic condition were not consistently more poorly recalled than the remaining conditions in that experiment. In the present experiment, we chose aircraft that were not, and never would be, traffic for any other aircraft; we called these Never-traffic. However, because of the necessity of knowing the two-dimensional position of *all* aircraft to make a determination of aircraft importance, the recall of the location of the Never-traffic aircraft should be comparable to that of the important aircraft.

A second motivation was to test how controllers remember altitude and ground speed. Do controllers remember the *exact* altitude and ground speed of an aircraft (i.e., the verbatim details) or remember this information in other ways? Research on cognitive development suggested that gist information, and not verbatim information, was crucial for reasoning (Brainerd & Reyna, 1993). Cognitive developmentalists discovered that verbatim memory for critical background information in a reasoning problem was independent of the quality of reasoning that resulted. For example, memory for the exact premises of a transitive inference problem (A > B, B > C) was unrelated to the likelihood of making the correct inference (A > C) (Brainerd & Kingma, 1984). Furthermore, this memory-independence effect continues into adulthood and appears to hold across a wide range of situations (e.g., attitude change, Hastie & Park, 1986; numerical reasoning, Klapp, Marshburn & Lester, 1983).

There are several memorial advantages to the encoding of gist over the encoding of verbatim details (Brainerd & Reyna, 1990; Reyna & Brainerd, 1992). These include stability, ease of retrieval, and ease of manipulation. Furthermore, there are several processing advantages to encoding the gist, including simplified processing, increased accuracy, and reduced effort. For example, two aircraft crossing at the same altitude is a problem, regardless of the altitude. Consequently, rather than encoding that AAL is at FL230 and SWA456 is at FL270, perhaps controllers encode only the "gist" (i.e., SWA is higher than AAL). A similar argument for the simplification of the controllers' mental load through "gist&cation" can be found in Moray (1990). Consequently, we explicitly tested for the use of a relational (gist-based) representation.

A final motivation for this experiment was to examine the impact on memory performance of one interpretation of the concept of free-flight (FAA, 1995). Currently, most aircraft fly along published routes through the sky. As a result of flying on these structured routes, there are particular points in a controller's airspace where routes intersect and merging aircraft may potentially conflict. However, when aircraft fly on direct or straight-line (unstructured) routes through the sector, aircraft may intersect at any point in the airspace. Because the FAA is planning a move to a free-flight environment (FAA, 1995), we thought it timely to include this manipulation in our study. The use of unstructured routing could increase the workload on the controller, which might lead to more simplification through gist processing and a greater reliance on a relational representation.

There are actually several stages between structured routing (the current system) and true free-flight (RTCA, 1995). These range from providing more optional routes to an aircraft to reduced separation standards requiring new technology in the cockpit and on the ground. We were unable to simulate true free-flight because the simulation facilities at our disposal would not support it. Consequently, we approximated free-flight as best we could by designing one scenario in which most of the aircraft followed unstructured routes.

Method

Participants

Eleven full-performance level (FPL) controllers participated. They had been full-performance level or FPL controllers for an average of 14.0 years. The controllers last worked in the field an average of 4.0 years before, with a range of .5 to 6.8 years. All participants were instructors at the FAA Academy and were familiar with the AeroCenter airspace used in the experiment. Participants worked alone and were responsible for all aspects of the sector, including communicating with pilots and other centers, and performing their own strip marking.

Materials

The experiment was conducted at the Radar Training Facility (RTF) of the Mike Monroney Aeronautical Center in Oklahoma City. Two 30-minute scenarios were created and modified with the help of the SME. One scenario used structured routing, in which most aircraft flew on standard routes. In the other scenario, most aircraft used unstructured routing, which we instantiated as straight-line routes from departure point to destination. The scenarios included a mean of 36 aircraft; 33 were overflights, 0 were arrivals, and 3 were departures. On average, there were 15 aircraft displayed at the time of testing. The order of the scenarios was counterbalanced across participants.

Procedure

There were four conditions. There were two Traffic conditions. One consisted of two aircraft converging at the same altitude (Traffic-level), and the other consisted of one aircraft climbing through the altitude of another (Traffic-climb). The Pre-traffic condition consisted of two aircraft that would intersect in the future. Finally, a Never-traffic condition was included as a more effective control. Aircraft importance was operationalized as temporal proximity to loss of separation. For the two Traffic conditions, the time window was approximately two minutes, the Pre-traffic aircraft were about 10 minutes apart at the time that the scenario was halted, and the pair of Never-traffic aircraft would never conflict with each other (or with any other aircraft within the time frame of our scenario). A given aircraft served in only one condition per scenario and was tested only once.

After studying and ordering the flight strips to their liking, the participants were told which scenario they were controlling (structured or unstructured routes). They were also told that all the aircraft on the PVD were under their control, and that none were currently in conflict.

At three points during each 30-minute scenario, the scenario was paused. We used a triggering event as a signal (e.g., a request from a particular pilot for an altitude, speed, or heading change); the signal occurred at approximate 1 O-minute intervals. The participant was turned away from the PVD and strip bay and completed two tasks.

The first task was Map recall. The computer presented a replica of the AeroCenter airspace (66% of its normal size) with no aircraft displayed. The participant used a trackball to move a cursor to the two-dimensional location of each aircraft on the radar and clicked to mark the location. Only after completion of the next task did we return to this task and ask the participant to indicate which call sign went with which aircraft.

After completing Map recall, the participant answered a series of questions about various aircraft. During this portion of the study, a full-size paper replica of the sector was provided with magnets designating all of the aircraft in position and their call signs. Three types of questions regarding the ground speed and altitude were asked about each pair of Traffic-level, Traffic-climb, Pre-traffic, and Never-traffic aircraft. The first question concerned relational information. For example, "is AAL123 higher, lower, or level with DAL246?" The second question concerned range information for each aircraft in the pair (ordered randomly). For example, "give an upper and lower bound for which you are 95% certain that the actual speed of AAL lies between them." The final question concerned the verbatim or exact data for each aircraft in the pair (ordered randomly). For example, "What is the exact altitude of AAL123?" We varied whether speed or altitude was asked about first. The order of the conditions was always Pre-traffic, Never-traffic, Traffic-climb, and Traffic-level.

A practice phase was completed prior to beginning the experiment that included a sample Map recall and an example of each question type. It began with the experimenter demonstrating how to respond to the various questions, and concluded with the participant completing the same practice trial. Participants completed two scenarios (randomly ordered) and received a 20-minute break between scenarios. After completion of both scenarios, a demographic questionnaire was administered.

Results and Discussion

Map Recall. In Map recall, locations corresponding to 79.6% of the aircraft were recalled. Recall varied slightly with condition and type of scenario producing a significant interaction ($F(3, 8) = 4.37$, $MSe = .094$) due to worse recall for the Traffic-climb aircraft in the structured scenario (95% vs. 74% respectively, $t(10) = 3.57$, $se = .052$). All $p <$ .05 unless otherwise indicated. However, this difference was due to very poor recall (27%) of a single aircraft in that condition, perhaps due to a computer error. Once this aircraft was removed, the interaction was eliminated ($F(3, 8) = 2.09$, $p = .18$, $MSe = .086$). In general, the two-dimensional positions of all the aircraft, even the Never-traffic aircraft, were equally well remembered.

We next computed the distance between the actual two-dimensional position of the aircraft and the recalled position of the aircraft (see Figure 1). We again excluded the previously mentioned Trafrc-climb aircraft. Controllers were quite accurate in recalling the position of the aircraft, only missing the actual position on the radar by an average of 2.85 cm (about 9.6 nautical miles). The mean missed distance varied significantly with condition ($F(3, 7) = 19.8$, $MSe = 1.4$ (one participant's data were lost due to computer malfunction). Post-hoc dependent t-tests collapsing across scenario type showed that the placement of Traffic-level aircraft

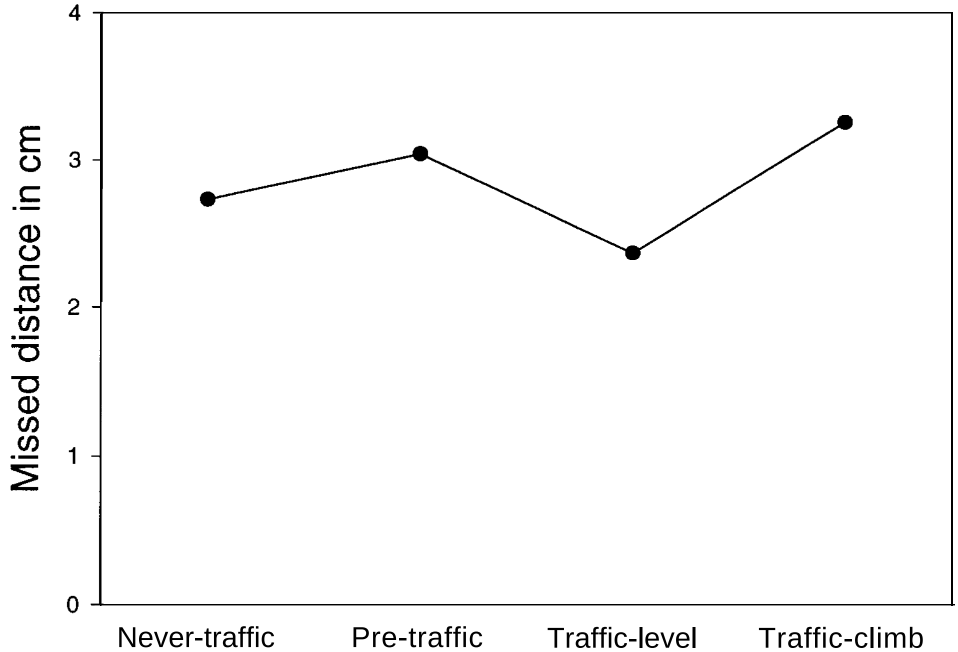

Figure 1. Missed distance in Map recall for the Never-traffic, Pre-traffic, Traffic-level, and Traffic-climb conditions.

was better than only the Pre-traffic condition *(t(10)* = 3.67, s_{diff} = .21), but only by .8 cm. Participants remembered, with a high degree of accuracy, the two-dimensional position of almost every aircraft on the PVD. Furthermore, this was largely unaffected by aircraft importance: The missed distance for the Traffk-level aircraft was only .4 cm better than the Never-traffic aircraft (not significant).

Question battery. A repeated-measures ANOVA was conducted on the responses to the "exact" questions. There was an interaction between scenario type (structured vs. unstructured) and type of question (altitude vs. speed, F(1, 9) = 11.29, MSe = .004); the interaction was the result of worse altitude recall in the unstructured scenario (structured = 63%, unstructured = 50%, *t(9) = 4.95,* s_{diff} = .024). At present, this finding could have been caused by any number of factors besides the use of unstructured routes, including our particular implementation of free-flight, or differences between the scenarios (e.g., complexity, specific conflicts to be resolved). Nevertheless, with the impending implementation of free-flight, this finding raises a cautionary flag and warrants further study.

There was a main effect of type of question, with recall of altitude superior to recall of speed (F(1, 9) = 11.61, MSe = .022, mean altitude = 56.5%, mean speed = 20.6%). There were no differences across conditions for the speed question (see top panel of Figure 2). Moreover, the only condition above chance for speed was the Pre-traffic condition (and only 3% above chance at that).

For altitude recall (see bottom panel of Figure 2), post-hoc tests (collapsed over scenario type, which did not differ with condition) showed that the Traffic-level and Pre-traffic aircraft were superior to the Traffic-climb and Never-traffic aircraft (minimum *t(9) = 3.63,* s_{diff} = .05). Our refined importance hypothesis expected that the recall of altitude for the Traffic-level and Pre-traffic aircraft (important aircraft) would be better than for the Never-traffic (unimportant) aircraft. However, we also expected that performance for the Traffic-climb aircraft would be very good because these aircraft were also important.

It turned out that the relatively poor performance for the Traffk-climb condition was somewhat misleading. In our instantiation of this conflict, one aircraft in the pair was level and the other was not. When we scored the accuracy of these two aircraft separately, we found that altitude recall for the level Traffic-climb aircraft was 72.2%, comparable

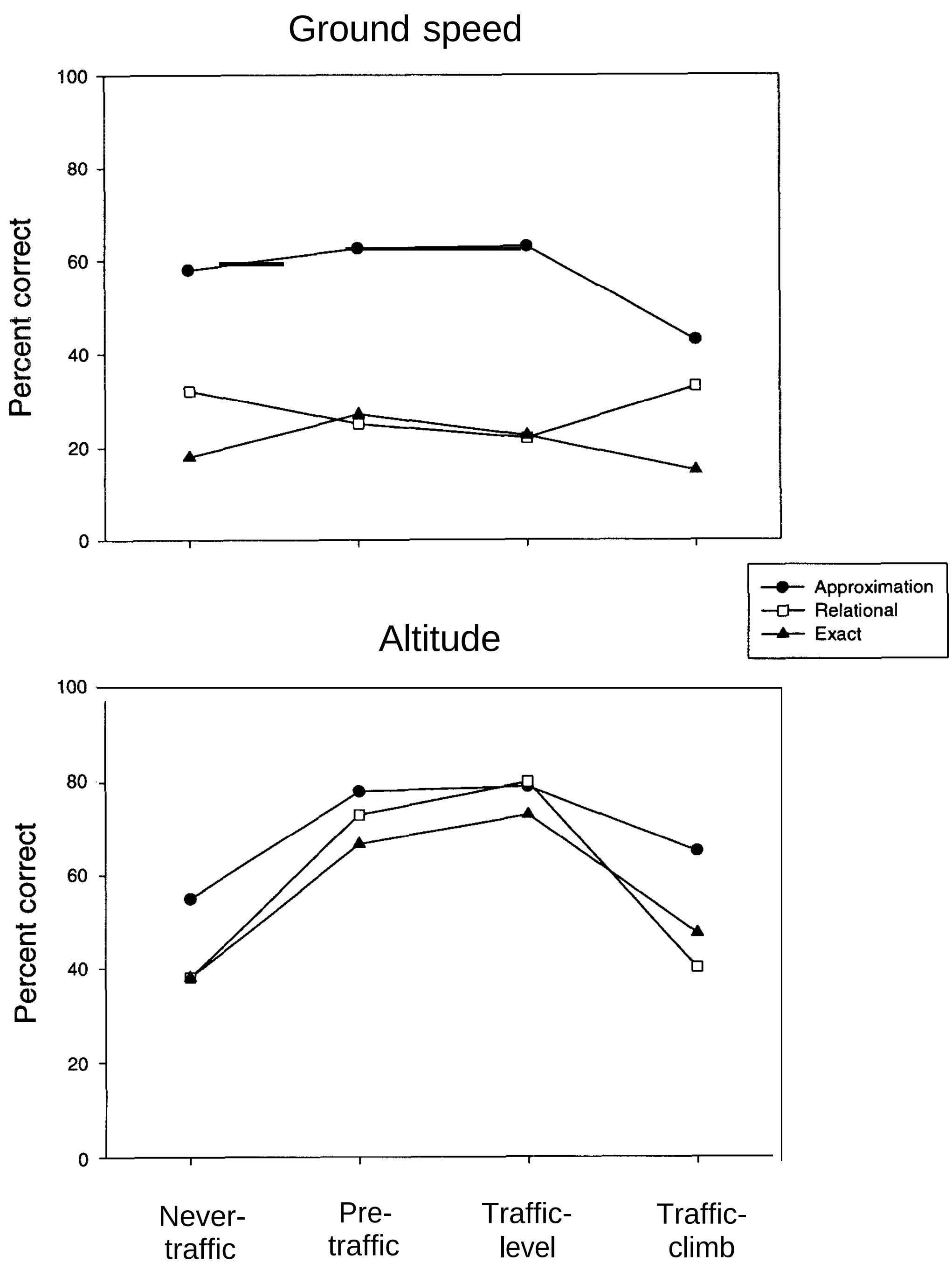

Figure 2. Percent correct for the Never-traffic, Pre-traffic, Traffic-level, and Traffic-climb conditions for exact, relational, and approximation scoring. The top panel is ground speed and the bottom panel is altitude. Data collapsed over scenario type.

to that of the Traffic-level and Pre-traffic aircraft. However, performance on the *climbing* Traffic-climb aircraft was very bad, only 18.3%. Our data indicated that the participants were unaware of this. This was supported by their answer to the approximation question The width of the participant-selected range did not differ with condition (altitude mean width = 2280 feet, speed mean width = 37.6 knots), which meant that the controllers did not compensate for their poor memory for the altitude of the climbing Traffic-climb aircraft by stretching their confidence range. In fact, the mean range of the climbing Traffic-climb aircraft was actually narrower than for the level Traffic-climb aircraft ($t(9) = 3.06$, $s_{diff} = 8.05$, mean climbing Traffic-climb = 1303 feet, mean level Traffic-climb = 3280 feet).

The poor performance of the climbing Traffic-climb aircraft might have been due to the phrasing of the questions. Although we asked for the assigned altitude of the climbing Traffic-climb aircraft, some participants might have tried to estimate its current altitude, mid-climb. However, performance did not improve to the level achieved by the level Traffic-climb aircraft (improving to only 30.3%) when we re-scored the data counting, as correct either the assigned altitude or ±2000 feet around the current altitude. Another possible reason for the poor performance related to the climbing Traffic-climb aircraft might be due to the design of the scenarios; these aircraft typically had just entered the airspace prior to the time that we stopped the scenario and had begun their climb just prior to the scenario being paused. However, this is an unlikely explanation; the late entry into the scenario was also the case for the Pre-traffic aircraft and altitude recall for these aircraft was very good.

Poor memory for the altitude (current or assigned) of aircraft that were climbing might signal poor SA. After all, 83% of the operational errors (loss of 5 miles horizontal and/or 1000 feet vertical separation) at en route facilities in 1993 involved aircraft changing altitudes (i.e., descending through an altitude occupied by a level or climbing flight (Durso and associates, 1996). On the other hand, it might be sufficient that the controller remembered that the aircraft was below, at, or above the aircraft with which it was in conflict. If so, we should see evidence that the altitude of the climbing Traffic-climb aircraft was remembered relationally.

Performance on the exact, relational, and approximate, speed and altitude questions were compared by subtracting out an estimate of chance. Chance was 1/3 for relational responses (three possible responses), and the SME estimated chance to be 1/10 for exact responses. Chance was difficult to compute for the approximation questions as it depended on the width of the participant-selected range. However, one estimate would be 50%-the answer either fell within the range (which, on average, covered about half of the possible range) or it did not. We subtracted out these estimates of chance and compared performance across exact, relational, and approximate questions. Performance on exact was best for altitude (exact = 46.5% vs. approximation = 19.2% and relational = 24.9%, all different from 0, exact greater than the other two, which did not differ). Controllers remembered the exact altitude of the aircraft, especially the important aircraft.

For speed, only exact was above chance (exact = 10.6% vs. approximation = -6.8% and relational = -5.0%, negative percentages indicated that performance was below chance, although not significantly). Contrary to Gronlund and associates' (1997) Experiment 2, speed was poorly remembered relationally in the present experiment. This might have been the result of the increased emphasis on speed in prior experiments, which all involved the sequencing of aircraft. Improved memory for flight data highlighted by the scenario was consistent with a second hypothesis proposed by Means and associates (1988)-that the type of control determined what flight data were remembered about an aircraft. For example, Means and associates found that vectoring an aircraft led to better retention of its routing information. If this hypothesis was true, speed was better remembered in Gronlund and associates' Experiment 2 because it was more important than in the present experiment. Despite the increased emphasis on speed in Gronlund and associates' Experiment 2, memory for exact speed was remembered equivalently in the two experiments, although not terribly well (about 12% better than chance). Interestingly, when speed was remembered, it was better remembered relationally than exactly.

A repeated-measures ANOVA was conducted separately on the relational and approximation questions. For the relational questions, the three-way interaction was significant ($F(3, 27) = 9.29$, MSe =

.075), as were two two-way interactions: question type interacted with scenario type (F(l) *9) = 7.76,* MSe = .06) and with condition (F(1, 9) = 22.35 MSe = .033). There was a main effect of type of question; altitude recall was superior to speed recall (F(1, 9) = 53.26, MSe = .068, mean altitude = 57.9%, mean speed = 27.9%), and a main effect of condition (F(1, 9) = 5.05, MSe = .053, shown in Figure 2). Post-hoc tests showed that there were no significant differences among conditions for speed, but for altitude, Traffic-level and Pre-traffic performance were superior to the Never-traffic condition (minimum t(9) = 5.16, s_{diff} = .07). We could not separately analyze the climbing Traffic-climb and the level Traffic-climb aircraft because the relational question necessarily considered both aircraft.

For the approximation question, a response was scored as correct if the actual altitude or speed fell within the participant-selected range. Question type interacted with condition (F(1, 9) = 3.34 MSe = .035). There was also a main effect of type of question; altitude recall was superior to speed recall (F(1, 9) = 7.11, MSe = .085, mean altitude = 65%, mean speed = 43.3%), and a main effect of condition (F(1, 9) = 7.95, MSe = .039, shown in Figure 2). Post-hoc tests showed no significant differences among conditions for speed. For altitude, Traffc-level and Pre-traffic performance were superior to the Never-traffic condition (minimum t(9) = 4.20, s_{diff} = .058). Performance on the level Traffic-climb aircraft was 80.4%, comparable to performance of the Traffic-level and Pre-traffic aircraft; performance on the climbing Traffic-climb aircraft was poor, only 45.4%.

Conclusions

Situation awareness is assumed to be central to successful air traffic control performance (e.g., Endsley, 1995a). The products of memory are viewed as central to achieving SA (Endsley, 1995b). What have we learned about the role of memory in air traffic control?

We have initial support for the hypothesis that controllers classify aircraft into two categories (important vs. unimportant) based on their knowledge of the two-dimensional position of the aircraft. Importance had minimal effect on memory for aircraft position, but the exact altitude (but not ground speed) of important aircraft was better recalled.

We refined the importance hypothesis first proposed by Gronlund and associates (1997) by showing that aircraft importance was a binary-valued construct: An aircraft either was not traffic for any other aircraft (unimportant), or it was or might be (important). This result would be consistent with Means and associates' (1988) "hot/cold" hypothesis if we assumed that important aircraft received more control actions. Nevertheless, we feel that the operative factor was the aircraft importance, not the amount of control.

The level of recall and the accuracy of the two-dimensional placement of the aircraft was excellent (see also Means and associates, 1988; Vortac and associates, 1993). The variables we examined (see also Gronlund and associates, 1997) did not differentially affect memory for this information. We think this is because it is important for the controllers to know the two-dimensional location of every aircraft on the PVD to make a determination of whether or not an aircraft is important.

Because controllers classify importance based on knowing an aircraft's two-dimensional position, they were able to more effectively distribute their mental effort. This resulted in better memory for the altitude of important aircraft. These findings can be used to improve the validity of recall measures of SA. It appears that there are some aircraft for which the controller should be expected to remember more, and other aircraft about which the controller could remember little. Not all aircraft are equally important to the controller, and measures of SA should not assume that they are. Of course, whether an asymmetric distribution of mental resources to important versus unimportant aircraft is an optimal strategy awaits improved performance measures that reflect moment-to-moment fluctuations. At this point, all we can be certain of is that the controllers do asymmetrically distribute their mental resources; we cannot say whether that is optimal. If it did turn out to be optimal, it would be important to determine if the recall level obtained for the flight data from unimportant aircraft reflected incidental encoding or a reduced level of normal encoding. If it reflected the latter, an improved interface or changes to training procedures could attempt to further redirect "wasted" resources away from the unimportant aircraft.

The controllers remembered altitude exactly, rather than approximately or relationally. This was surprising given the information reduction advantages of remembering something relationally (Brainerd & Reyna, 1990; Moray, 1990). Although there are other ways that gist information might be represented for altitude (e.g., is there any other aircraft at the same altitude as AAL123?), we believe that exact altitude reflects the way this information was remembered in these scenarios. Whether this is the most efficient or effective way to remember altitude information is not known.

In contrast to altitude, exact ground speed was poorly remembered (see also Means and associates, 1988). However, the increased relevance of speed in the sequencing problems of Gronlund and associates (1997, Experiment 2) was sufficient to produce a fairly high level of performance for relational speed. This provides some support for the Means and associates (1988) hypothesis that the type of control affected what was remembered. Refined measures of SA should take into account the relevance of the flight data to the particular scenario being tested. For example, if additional empirical work establishes that ground speed is remembered relationally (when it is remembered), a controller's SA would be most accurately measured by tapping ground speed in that way.

It is still possible that we inappropriately tapped the controllers' memory for speed. We have not yet queried speed in terms of time. In other words, a controller might not remember which of two aircraft is faster, nor by how much, but they might remember that the two aircraft will lose separation in j minutes if something is not done. To make that determination would require memory for aircraft speed, although not in the way we have examined it.

One finding that warrants further study was the very poor memory for the altitude of climbing traffic, especially given the large percentage of operational errors involving aircraft that are changing altitude (Durso et al., 1996). At present, it is unclear to what extent this is a general problem versus something unique to our scenarios and methodology. If our results are indicative of a general problem, monitoring the altitude of climbing (and presumably descending aircraft) is a situation in which the controllers would benefit from extra assistance. Perhaps some kind of gist representation of the relevant information might help to alleviate the problem. For example, climbing aircraft could be color-coded to signal their altitude relative to another aircraft or relative to various fixed altitudes.

A second finding that warrants additional study was that exact altitude was remembered relatively poorly for unstructured scenarios. With the imminent implementation of free-flight, this raises some concerns. Is this finding the result of attention being drawn away from altitude information due to the increased difficulty of detecting conflicts at all points in a sector rather than at designated intersection points? A cognitive aid that facilitates the identification of potential conflicts might ease the transition to a free-flight environment by freeing up mental resources that could then be redirected to altitude information.

The cognitive abilities of air traffic controllers will be further stretched by the advent of free-flight and the ever-increasing volume of air traffic. The continuing investigation of how controllers perform their jobs and memory's role in support of that job can continue to be profitably applied to the development of cognitive aids and the re-design of interfaces. These modifications will allow the controller to better manage the complex, dynamic air traffic control system of tomorrow.

References

Adams, M.J., Tenney, Y.J., & Pew, R.W. (1995). Situation awareness and the cognitive management of complex systems. Special Issue: Situation awareness. *Human Factors, 37,* 85-104.

Anderson, J.R. (1995). *Cognitive psychology and its implications,* 4th edition, New York: W.H. Freeman and Company.

Brainerd, C.J., & Kingma, J. (1984). Do children have to remember to reason? A fuzzy-trace theory of transitivity development. *Developmental Review,* 4, 311-77.

Brainerd, C.J., & Reyna, V.F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review, 10, 3-47.*

Brainerd, C.J., & Reyna, V.F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review, 100, 42-67.*

Buckley, E.P., DeBaryshe, B.D., Hitchner, N. & Kohn, P. (1983). Methods and measurements in real-time air traffic control system simulation (Report No. DOT/FAA/CT-83-26). Atlantic City, NJ: Federal Aviation Administration Technical Center.

Durso, F.T., Truitt, T.R., Hackworth, CA., Ohrt, D.D., Hamic, J.M., & Manning, C.A. (1996). Factors characterizing en route operational errors: Do they tell us anything about situation awareness? In D.J. Garland, & M.R. Endsley (eds.), *Situation Awareness,* 189-196.

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

Endsley, M.R. (1995a). Toward a theory of situation awareness in dynamic systems. *Human Factors,* 37, 32-64.

Endsley, M.R. (1995b). Measurement of situation awareness in dynamic systems. *Human Factors,* 37, 65-84.

Federal Aviation Administration (1995). *Air Traffic Service Plan:* 1995-2000, Washington, DC.

Gronlund, S.D., Dougherty, M.R.P., Ohrt, D.D., Thomson, G.L., Bleckley, K., Bain, D., Arnell, F., & Manning, CA. (1997). The role of memory in air traffic control. (Report No. DOT/FAA/AM-97122). Washington, DC: Federal Aviation Administration Office of Aviation Medicine.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93, 258-68.*

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, 240-63.*

Marshak, W.P., Kuperman, G., Ramsay, E.G., & Wilson, D. (1987). Situational awareness in map displays. In *Proceedings of the Human Factors Society 31st Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

Means, B., Mumaw, R.J., Roth, C., Schlager, M.S., McWilliams, E., Gagné, E., Rice, V., Rosenthal, D., & Heon, S. (1988). *ATC training analysis study: Design of the next-generation ATC training system*. (Report No. FAA/OPM 342-036) Washington, DC: Department of Transportation/Federal Aviation Administration.

Moray, N. (1990). A lattice theory approach to the structure of mental models. *Philosophical Transactions of the Royal Society of London, B327, 577-83.*

RTCA (1995). *Final report of the RTCA Task Force 3: Free flight implementation,* Washington, DC.

Reyna, V.F., & Brainerd, C.J. (1992). A fuzzy-trace theory of reasoning and remembering: Paradoxes, patterns, and parallelism. In A.F. Healy, S.Kosslyn, & R.M. Shiffrin (eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (pp. 235-60). Hillsdale, NJ: Erlbaum.

Vortac, O.U., Edwards, M.B., Fuller, D.K., & Manning, C.A. (1993). Automation and cognition in air traffic control: An empirical investigation. *Applied Cognitive Psychology 7, 631-51.*