2003

Time vs. certainty : pilot preference and conflict alerting in free flight

Paul Picciano
San Jose State University

TIME VS. CERTAINY: PILOT PREFERENCE AND CONFLICT ALERTING IN FREE FLIGHT

A Thesis

Presented to

The Faculty of the Graduate Program in Human Factors and Ergonomics

San Jose State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Paul Picciano

Adviser: Dr. Kevin Corker

May 2003

UMI Number: 1415729

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

APPROVED FOR THE GRADUATE PROGRAM
IN HUMAN FACTORS AND ERGONOMICS

(Dr. Kevin Corker – Thesis Chairperson signs here)

(Dr. Kevin Jordan – Thesis Committee Member signs here)

(Sandy Lozito– Thesis Committee Member signs here)

APPROVED FOR THE UNIVERSITY

ABSTRACT

TIME VS. CERTAINY: PILOT PREFERENCE AND CONFLICT ALERTING IN FREE FLIGHT

by Paul Picciano

There are two important but conflicting parameters in airborne conflict alerting. The first is look-ahead time (i.e., time available prior to conflict). The second parameter, certainty, is the probability that a loss of separation is inevitable. The present study was conducted to identify pilot preference concerning the irreconcilable trade-off between the two. Pilots were required to assess traffic situations and decide if a maneuver was required to avoid a conflict. The pilots were assisted by automation that provided alerts with three different look-ahead times (2, 4, and 8 minutes) and two certainty level thresholds (high = 99%, low = 75%). The results indicated that only the look-ahead time variation impacted pilot behavior. No significant differences were observed between the high and low certainty conditions.

DEDICATION

I would like to thank Sandy, Kevin, and Kevin for presenting me with such wonderful opportunities at SJSU/NASA Ames while pursuing my degree. I truly appreciate the wisdom and camaraderie shared along the way.

ACKNOWLEDGEMENTS

This experiment would not have been possible without the programming expertise of Dominic Wong. Thanks also to Steve Sheldon, Walt Johnson, and Vern Battiste for making the traffic display (CDTI) available.

I would also like to thank my co-workers at NASA-Ames for their input, assistance, and support throughout this process. Thanks to San Jose State University Foundation and NASA-Ames Research Center for providing facilities and equipment.

This work was funded through a grant courtesy of the Federal Aviation Administration (FAA) from the office of the Chief Scientist for Human Factors. Also, the Advanced Air Transportation Technologies (AATT) Project, part of the Aviation Systems Program, helped support this work.

Table of Contents

List of Tables....viii
List of Figures....ix
1. Introduction....1
    Background....3
        Free Flight....3
        Certainty....3
    Hypotheses....11
2. Method....18
    Participants....18
    Apparatus....18
    Traffic Display and Alerting....18
    Procedure....19
    Independent Variables....22
    Dependent Measures....23
3. Results....24
4. Discussion....28
5. Conclusion....32
References....34
Appendix A   Post-Run Questionnaire....38
Appendix B   ANOVA Tables....40
Appendix C   Graphical Representation of the Means for Each Dependent Measure....44
Appendix D   Conflicts and Alerting Algorithms....51
Appendix E   Uncertainty Due to Winds and Navigation....56
Appendix F   Scenario Construction....60
Appendix G   Traffic Display....64
Appendix H   Free Flight Brief....67
Appendix I   Participant Profile....69
Appendix J   Consent Form....71

List of Tables

Table 1 Mean pilot maneuver times (and standard deviations) for each dependent measure relative to 1) start of scenario - *0Time* 2) the alert -*AlertTime* 3) the loss of separation event –*LOSTime*......................24

Table 2 Subjective responses for each participant from post-run questionnaire....27

Table 3 ANOVA output for dependent measure 0Time....................................... 41

Table 4 ANOVA output for dependent measure AlertTime 42

Table 5 ANOVA output for dependent measure LOSTime 43

List of Figures

Figure 1 Cockpit Display of Traffic Information (CDTI) in an alert state............. 20

Figure 2 Maneuver App. Used in the experiment to collect time of maneuver, type of maneuver, and intruder identification..21

Figure 3 Means and standard error bars for maneuver times shown for each dependent variable (0Time, AlertTime, and LOSTime) by each look-ahead time level..26

Figure 4 Means and standard error bars for elapsed times from scenario start at which pilots initiated resolution maneuvers for each look-ahead time level 45

Figure 5 Means and standard error bars for maneuver times relative to the alert for each look-ahead time level (a negative value indicates the maneuver occurred before an alert was issued)...46

Figure 6 Means and standard error bars for number of seconds prior to the loss of separation event the maneuver was initiated for each look-ahead time level...47

Figure 7 Means and standard error bars for elapsed time from scenario start at which pilots initiated resolution maneuvers for each level of alert certainty. 48

Figure 8 Means and standard error bars for maneuver time relative to the alert for each certainty level..49

Figure 9 Means and standard error bars for number of seconds prior to the loss of separation event the maneuver was initiated for each certainty level...50

1. INTRODUCTION

How will pilots self-separate from neighboring aircraft and resolve future conflicts? In order to support a more distributed air traffic management system (i.e., transfer some self-separation and collision avoidance responsibility from ground controllers to the flight deck), pilots will need information about proximal aircraft. In addition to the current traffic configuration, future position data for potentially conflicting aircraft will also be valuable for preserving conflict-free routes. Further, an alerting system will likely be useful in aiding pilots in these conflict avoidance tasks (Hoekstra, van Gent, & Ruigrok, 1998).

There are several decisions to be made concerning the parameters of the conflict alert, and the optimal alerting strategy is not obvious (Johnson, Battiste, & Bochow, 1999; Paielli & Erzberger, 1997). Due to uncertainties in the prediction of aircraft trajectory, a conflict warning can not achieve 100% certainty (Gempler & Wickens, 1998; Kuchar, 2001; Paielli & Erzberger, 1997;Yang & Kuchar 1997). A conflict warning is an alert indicating two aircraft will be closer than a prescribed safety minimum. This is known as a loss of separation (LOS). Thus, there is a fundamental trade-off between the amount of time available before a conflict occurs (i.e., look-ahead time), and the probability that the conflict will actually occur (i.e., certainty of conflict).

Accordingly, conflict alerting systems on the flight deck can adopt different strategies based on this trade-off. One design approach could be based on expanding look-ahead times. Such a system would alert the pilot to a potential conflict situation well in advance of the critical event (e.g., a loss of separation). If notifications at greater times and distances are issued, smaller course deviations can resolve the conflict compared to maneuvers required at shorter look-ahead times (and closer distances).

The problem with early notification is that the reliability (i.e., certainty) of the prediction is lower than the reliability of a prediction for the same conflict made at a later time. Thus, there is a chance of producing a false alarm. A false alarm is an alert that was issued when there would have been no conflict and no reason to maneuver. False alarms have been shown to have a negative impact on an operator's trust and use of a system (Parasuraman, 1997).

Conversely, an alerting system could be designed around a certainty threshold. For example, an alert would be suppressed until the alerting logic calculated the conflict probability to be greater than a given value of certainty (the threshold). This would reduce the likelihood of generating a false alarm. But, if the threshold was set too high, it may leave very little time or space to make an avoidance maneuver because certainty increases are gained at the expense of look-ahead time. The advantage with this method is the reduction of false alarms, but this introduces the possibility of late and missed alerts. An alert is considered late when the information it provides is no longer useful in preventing the incident it was intended to warn against.

To summarize the problem statement, it is not possible for a pilot to have the best information (highest certainty) <u>and</u> the most time to make decisions and initiate maneuvers (look-ahead time). Therefore, a pilot will be forced to make a decision based on this trade-off. The purpose of this study is to examine pilot preference concerning this choice. It is hoped this research can begin to identify what information the human operator values when challenged with self-separation and conflict avoidance tasks.

Background

Free Flight

An air traffic management concept in which pilots may perform such tasks as self-separation and conflict avoidance is currently referred to as "free flight." Free Flight was formally introduced in 1995 and defined by RTCA Task Force 3 (RTCA, 1995).

> Free flight is a safe and efficient operation capability under instrument flight rules (IFR) in which the operators have the freedom to select their path and speed in real time. Air traffic restrictions are only imposed to ensure separation, to preclude exceeding airport capacity, to prevent unauthorized flight through Special Use Airspace (SUA), and to ensure safety of flight. Restrictions are limited in extent and duration to correct the identified problem. Any activity which removes restrictions represents a move toward free flight. (p. 23)

Anticipating the benefits of an expanded free flight concept (RTCA, 1995):

> Free flight can provide the needed flexibility and capacity for the foreseeable future. As its basis, the concept enables user preferred (dynamic) flight paths for all airspace users through the application of CNS/ATM technologies and the establishment of ATM procedures that maximize flexibility while assuring positive separation of aircraft. (p. 25)

In contrast to the present architecture, "The primary difference between today's direct route clearance and free flight will be the pilot's ability to operate the flight without specific route, speed, or altitude clearances." (RTCA, 1995 p. 25).

Certainty

In providing traffic information and conflict alerts to support free flight, there is intrinsic uncertainty in the measurements and calculations. These uncertainties will impact the information presented on traffic displays in the cockpit as well as conflict alerts presented to pilots. Paielli and Erzberger (1997), concede "errors in prediction are unavoidable," concerning the trajectories of aircraft. (p. 1). Uncertainty is the result of many factors. Measuring devices and navigational systems are limited in their accuracy. Even with the deployment of advanced technology such as the global positioning system (GPS), determining aircraft position is imperfect. The accuracy of GPS is limited to locating an aircraft within 13 meters horizontally and 22 meters vertically with reliability to 95% (Stenbit, 2001).

Winds and weather also introduce error in predicting the future position of aircraft (Cole, Green, Jardin, Schwartz, & Benjamin, 2000; Gempler & Wickens, 1999; Paielli & Erzberger, 1997; Williams &Green, 1998). Cole, Green, Jardin, Schwartz, and Benjamin (2000) reported on wind prediction inaccuracies based on flight tests in 1992 and 1994. "Flight tests have indicated that wind prediction errors may represent the largest source of trajectory prediction error" (p. 2). Discrepancies of 20 knots were typically found between the predicted wind values and the actual wind values. A headwind calculation 15 knots in error projected over a 20 minute flight segment is sufficient to create an error of 5 nmi.

In developing self-separation and conflict avoidance tools, several researchers have encountered the difficulties of prediction error (Hoekstra et al., 1998; Johnson et al., 1999; Yang & Kuchar, 1997). A number of researchers have specifically noted the trade-off involved. Paielli and Erzberger (1997) declare, "The optimal time to initiate a conflict resolution maneuver is a trade off between time to conflict and certainty," (p.1). Johnson, Battiste, Delzell, Holland, Belcher and Jordan (1997) note the same conundrum:

> On the one hand, early predictions could lead to unnecessary maneuvering as crews respond earlier in an attempt to avoid low probability potential encounters. On the other hand, early maneuvering can be more efficient than late maneuvering with the crew being able to select more gradual maneuvers when given more time and distance to the potential conflict. (p.2)

Yang and Kuchar (1997) in developing a prototype conflict alerting logic also note:

> Because flow efficiency is a driver for free flight, it is desirable that conflicts be resolved using minor course, speed, or altitude changes well before emergency avoidance maneuvers are needed. It is also desirable, given the large number of aircraft in the air, that conflict alerts are only generated when necessary. However, the large amount of uncertainty in the free flight environment makes it difficult to determine how likely a projected conflict is to occur. The result is a tradeoff between alerting early to provide a large safety margin (and also producing unnecessary alerts) vs. alerting late to reduce unnecessary alerts (but requiring more aggressive avoidance maneuvers). (p. 768)

Though uncertainty in conflict prediction has been identified by various researchers, there still remains a significant amount of examination to be done. Few investigations have been made to directly assess the impact of uncertain alerting data on the pilot in the context of free flight.

One relevant study was performed by Gempler and Wickens (1998). In the experiment a traffic display was used to assist pilots in conflict avoidance. But instead of addressing the uncertainty of a conflict alert, they examined the uncertainty in predictor lines. A predictor line is a projection from each aircraft on the traffic display depicting where that particular aircraft will be at a future time. Most traffic displays that employ this feature simply use a narrow line extending from the apex of the aircraft symbol (Hoeksta et al., 1998; Johnson et al., 1999). Gempler and Wickens (1998) decided to display the uncertainty of this prediction by using diverging lines. Two lines extended from the nose of the aircraft to form a triangular shape. This figure represented the probability distribution of where the aircraft could be at a future time. Gempler and Wickens (1998) hoped that the display of uncertainty would help mitigate the effects of an invalid predictor. That is, it was believed the displayed uncertainty could assuage the consequences of false alert trials. No significant effects were found however, and they were unable to calibrate pilot trust in their system as they hoped to do. In their experiment, knowledge of certainty produced no observable effect on pilot behavior.

Other studies have also quantitatively probed reliability in hopes of correlating certainty levels with human sentiment. Examples have only been found beyond conflict alerting on the flight deck. In one such experiment, drivers were offered traffic reports with varying levels of certainty. The second study simulated an enemy engagement scenario where fighter pilots were tasked with making shoot/no shoot decisions. Finally, an experiment similar to the present study was used to examine the look-ahead time and certainty of conflict alerting in an air traffic control task. These will be discussed further.

A probe of acceptable certainty ranges concerning automobile traffic was performed by Kantowitz, Hanowski, and Kantowitz (1997). The researchers used a driving task to examine three levels of certainty and their effect on driver behavior and opinion. The goal was to discover an acceptable fidelity level for traffic information provided by the Advanced Traveler Information System (ATIS). Forty-eight participants were placed in a driving simulator and given a driving task in which real time information was provided to assist the motorist. Certainty levels of 100%, 71%, and 43% were utilized for the study.

The participants were charged with driving to a predetermined destination in the shortest time. To accomplish this, drivers had to select the links (i.e., road segments) with the least traffic and shortest travel time. In addition to the certainty levels, participants were placed in familiar and unfamiliar settings (a familiar or unfamiliar city). This variable enabled exploration of the link between trust in the automation and self-confidence of the operator. A system of penalties and rewards was constructed to provide incentive to use the automation. Any link chosen from the set comprising the optimal route was rewarded while drivers were penalized for selecting roadways with congestion. The participants were provided with feedback as to the success of their decisions after the run was complete (though real-time feedback was accrued by the observer as well, heavy traffic for example).

The between-subjects design assigned each of the two groups different certainty levels. All participants were run through four trials. Both groups had perfectly reliable information (100%) on their first two trials. Trials three and four were manipulated differently for each group. One group experienced 71% accuracy for the last two trials and the other group was given information with a reliability of only 43%.

It is not surprising that participants in both groups performed best when receiving 100% accurate information from ATIS. This is to be expected for both familiar and unfamiliar settings.

The findings of interest reside in the lower certainty levels. The participants provided with 71% accuracy for the second half of their trials reported that the information was still reliable and more importantly still useful. However, the group with 43% accuracy showed significantly worse performance and expressed frustration with the data.

Moving to a military domain, Banbury, Selcon, Endsley, Gorton, and Tatlock (1998) examined the reliability necessary for a fighter pilot to determine if an aircraft was friend or foe requiring a shoot/no shoot decision. Again, automation assisted the operators in performing this identification task. The alert system reported five levels of reliability: 97%, 94%, 91%, 79%, and 61%. The uncertainty complements to these values were also used as a presentation variable (3%, 6%, 9%, 21%, and 39%). The addition of the uncertainty values was used to study framing effects which, while not the focus of the current investigation, provided some interesting findings. For example, a delay in reaction time emerged when participants were presented with uncertainty values instead of confidence values. Banbury et al. (1998) suggested this lag was due to increased processing time; the participants converted uncertainty to confidence values (i.e., 3% uncertainty was transformed to 97% confidence in the mind of the pilot).

Significant results were obtained for the shoot/no shoot decision providing a range of certainty pilots were willing to accept. Independent of the framing issues, pilots chose to shoot significantly more times when confidence was 91% or greater (9% uncertainty or less), and refrained from firing when confidence was 79% and below. The values obtained from the intermediate certainty ranges did not significantly affect the decision to engage.

However, the reaction time data over these mid-certainty ranges is quite informative. The certainty condition of 97% elicited reaction times that were significantly faster than the 94% and 91% conditions. The decision to shoot was made quickly in this case. Interestingly, the reaction times in the 97% condition were not significantly faster than those in the 79% and 61% certainty treatments, which were overwhelmingly no shoot decisions. The authors suggest that pilots labored over their decision when the certainty range approached their apex of risk acceptability. The decision not to shoot with low confidence was made with rapidity similar to the shoot decision in the highest confidence condition.

Apparently, the high uncertainty (inherent in the low confidence values), created far too great a risk of fratricide and there was no need to contemplate the action. This implied pilots had comparable decision loads in deciding to abstain from firing in the two lowest certainty conditions, as they did making the decision to engage under greatest certainty.

This study suggests that the reliability of a decision support tool in a life-critical task must be of an appreciable fidelity in order to be utilized by the operator. Automation that reported data with greatest accuracy fostered superior performance. The critical issue is to discover at what point a decision support system becomes devalued. In the case of fighter pilots deciding to engage an enemy, it seemed that 97% accuracy was sufficiently close to a perfect system to provide data that was highly revered by the operator. But, this dropped quickly as 94% and 91% confidence levels did not produce results comparable to the 97% condition. In fact, workload seemed to increase as the automation provided information close to the highest level of risk a pilot was willing to tolerate.

The anxiety of the kill decision was again allayed when the certainty provided by the support tool fell below a pilot's comfort threshold. This information was unreliable to the extent that the pilot quickly concluded the risk of engagement was too great. With insufficient assurance, the pilots decided not to engage.

In the third experiment, Jürgensohn, Park, Sheridan, & Meyer (2001) performed an investigation in an air traffic control context to evaluate the effect of different alerting parameters on the operator's ability to detect conflicts. The look-ahead time of the alert as well as the reliability were manipulated in their experiment. Alert certainty levels of 100% reliability, 90% and 70% were utilized. The fourth certainty condition provided no alert to the operator and therefore no associated certainty. Paired with the reliability parameter were two warning time intervals. The operator was warned 5 or 15 seconds prior to the conflict event.

Their investigation of attention (using eye-tracking equipment), and response time revealed that behavior with less reliable alerting (70% condition), yielded results similar to the condition in which there was no alert. This result was obtained in the longer alerting time (15 seconds). For the warning time of 15 seconds, there was a noticeable effect of reliability on the operators' decision to act. The average response time was over 10 seconds for the alert that was 100% reliable while the 90% and 70% conditions were only between 5 and 10 seconds in advance of the conflict. The 70% condition was much closer to 5 seconds than was the 90% condition, prompting the researchers to compare this low level of certainty to the scenarios in which no conflict alert was provided (i.e., a missed alert). A near constant average reaction time was found when the alert gave only a 5 second warning (e.g., look-ahead time). The different reliabilities of the alert seemed to have no effect when the alert was administered with this brief look-ahead time.

From the aforementioned research examples, it can be seen that different levels of certainty are useful with different tasks. Attempting to identify a single certainty threshold suitable for all tasks and operators would be feckless. The individual differences connate in each operator will play a major role (Gempler & Wickens, 1999). Further, the criticality of the task could impact individuals' idiosyncrasies. The results of the previous experiments demonstrate this quite clearly. For a life and death decision Banbury et al. (1998) found performance suffered when certainty fell below 97%, while Kantowitz et al. (1997) found that a system reliability of only 71% could still prove "acceptable and useful" (p.173).

Hypotheses

In the dichotomy of longer alerting time and higher alerting certainty, it is hypothesized that one of the two may be dominant over the other under different conditions. Prior research has documented these effects of time and certainty (Banbury et al., 1998; Bliss, Gilson, & Deaton, 1995; Jürgensohn et al., 2001).

From a human performance perspective, it is postulated that some situations would elicit decisions based largely on the look-ahead time available. For example, with only a few seconds left before a critical event, it is unlikely an operator will investigate the certainty of the prediction before taking corrective action. This was the finding in the shorter alerting condition in the Jürgensohn et al. (2001) experiment. There is also theoretical support for this behavior.

Maule and Edland (1997) assert that decision makers under time pressure are forced to abandon some of the available information. They refer to this phenomenon as "selectivity." The time constraint does not permit all the information to be marshaled and the decision must be made with fewer data. When pilots are forced to make decisions with little look-ahead time (2 minutes), it is predicted that the certainty information will be ignored in the current experiment. Look-ahead time is believed to be the critical factor in motivating operator behavior in these cases. This would be akin to the results found by Jürgensohn et al. (2001).

A second instance in which look-ahead time may supercede the certainty of a conflict is at very large look-ahead times (8 minutes). A response latency is expected to emerge in the early stages of conflict. This resulting delay is anticipated regardless of the decision by the pilot. Based on a theoretical framework provided by Dror, Busemeyer, and Basola (1999), response "congruency" will impact the reaction time of a decision. This will be explained further.

At 8 minutes prior to a loss of separation, the risk level is relatively low. Dror et al. (1999) expect decision makers to assume more risk in lower risk situations and act more conservatively when faced with higher risk. Behavior consistent with this thinking produces "congruent responses" (p. 714). As applied to this thesis, accepting more risk in the simulation would result in maintaining course for some time after the alert. If a maneuver response is ultimately made, there will be a delay between the onset of the alert and the time of the maneuver.

Further, the model proposed by Dror et al. (1999) also predicts a delay in the case the pilot decides to maneuver earlier. They purport that "incongruent responses" (i.e., decisions that are not consistent with the level of risk), take longer than congruent responses. Therefore, it is predicted that even if pilots decide to maneuver immediately in the 8 minute condition a delay will emerge because it is an incongruent response. It is incongruent because the risk level is relatively low and an avoidance maneuver is a risk-avoiding decision.

Thus, either decision after an alert in the 8 minute condition is predicted to occur with some delay. Therefore, in the early stages of conflict it is believed the extended look-ahead time will produce a delay in the initiation of a resolution maneuver independent of certainty.

In contrast, it is also anticipated that situations exist such that the certainty of the alert is held in higher regard than the look-ahead time (Bandury et al., 1998; Bliss et al., 1995). This is hypothesized to be in a middle look-ahead time range where higher certainty could support earlier decisions to act. Here, the absence of time pressure (unlike the 2-minute condition), could permit the pilot to collect and process all available data (Maule & Edland, 1997). The certainty of the alert can be evaluated and incorporated into the decision. A time frame may exist in which an operator will only engage with sufficient certainty.

In order to investigate these theories, numbers related to look-ahead times and certainty levels were required to be used in the experiment. Little theoretical basis was available to arrive at appropriate values for practical use. Therefore, previous results found in the literature as discussed above were used to guide these decisions. Three look-ahead time values were chosen. The 2-minute condition, chosen as the late stage of conflict (i.e., the shortest amount of time available for a maneuver decision), was based on the work of RTCA Sub Committee 186, Working Group 1. (The RTCA serves in an advisory capacity to help craft policies and procedures for the FAA.) The choices of 4 minutes for the middle stage of conflict, and 8 minutes as the early stage (i.e., the greatest amount of time available for a maneuver decision), were derived from the work of Johnson et al. (1997).

The certainty parameters were also derived from previous experiments. In the current experiment 99% was used as the high level certainty threshold (99% was chosen instead of 100% to demonstrate the fallibility of the alerting logic). The low level threshold of certainty was balanced to provide a substantial departure from the high certainty condition while still proving useful. Based on the work of Banbury et al. (1998), Jürgensohn et al. (2001), and Kantowitz et al. (1997), 75% was chosen as the low certainty threshold for this experiment.

To map these fundamental ideas to the work covered by this thesis, the following hypotheses are put forth:

I. In the late stages of conflict resolution (e.g., 2 minutes/16nmi or less to the critical event), LOOK-AHEAD TIME is dominant over the certainty value.

To investigate the late stages of conflict resolution, a 2-minute look-ahead time condition was used. With such a short look-ahead it was hypothesized that certainty would be neglected in the operator's decision process and the response times (i.e., initiation of a maneuver), would be similar in the high and low certainty conditions. This was the result obtained by Jürgensohn et al. (2001). In the present study, it was also possible that the pilot could choose to maneuver *before* an alert was issued due to path and proximity information of aircraft on the traffic display.

II. There is a middle stage (4-6 minutes/40nmi prior to a critical event), of conflict resolution in which CERTAINTY is the dominant cue (that is the accuracy of the prediction is more important than the timing of the alert).

The middle stages of conflict resolution were considered by the researcher to be the 3-5 minute range and therefore a 4-minute look-ahead time was chosen based on previous research (Johnson et al., 1999). Here it was predicted certainty could influence the pilot's decision to act. A 4 minute warning with 99% certainty may elicit an immediate reaction, while the same look-ahead time with only 75% confidence may not produce as rapid a response. It was expected that certainty would play a role in these scenarios because this seemed to be on the perimeter of the comfort zone found by Johnson et al. (1999). By reducing the possibility of false alarms (with higher certainty), the benefits of early maneuvering could likely be realized (e.g., minimal deviation), since the risk of unnecessary maneuvers was reduced. This would be similar to the finding of Banbury et al. (1998) in which fighter pilots were slower to react when the certainly level dropped from 97% to 91%. Jürgensohn et al. (2001) demonstrated that lower certainty values resulted in a delayed response.

III. In the early stages (8 minutes/ 70nmi and greater in advance of the critical event), of conflict resolution, LOOK-AHEAD TIME is again predicted to be the dominant factor (there is so much time available that even almost perfect certainty would not elicit immediate action).

Finally, it was hypothesized that there exists an extended look-ahead time, that despite very high certainty, immediate action would be unlikely. This was tested in the condition in which an alert with 99% certainty was issued 8 minutes prior to a loss of separation. It's possible at extended look-ahead times the cost of an unnecessary maneuver would be greater than any benefit an immediate maneuver could provide. Waiting some time after an 8-minute alert does not significantly reduce maneuver options (Yang & Kuchar 1997). In addition, Dror et al. (1999) predict that maneuvers in this time frame were subject to delay because they would not be consistent with the level of risk. Therefore, immediate decisions to maneuver were not expected to be observed in this condition.

A prediction of these hypotheses is that an interaction of look-ahead time and alert certainty will emerge. In the early and late stages of conflict, it is believed that the time at which avoidance maneuvers are made will be similar in both certainty conditions. It is only the middle stage in which the reaction to maneuver should be faster in the high certainty condition than in the low.

2. METHOD

Participants

Twelve commercial airline pilots with glass cockpit experience were recruited for this experiment. This group of pilots had an average of over 14,500 (SD = 6135) total flight hours, averaging more than 5,500 hours with a glass cockpit. The age ranged from 38 to 59 years with a mean of 51.6 years. All but two had at least a limited knowledge of free flight.

Apparatus

Two Dell Dimension 8200 Series, Pentium 4 at 2.2 GHz desktop computers were used to administer the study. One computer was equipped with a 20 inch Dell 2000FP flat panel monitor and was used for the traffic display system. The second monitor was a 21 inch Dell P1130 Trinitron CRT. The data collection application was presented on this second monitor. Each Dell Dimension 8200 was stocked with an 80GB hard drive, 512MB RDRAM, and a 64MB NVIDIA GForce 3Ti 200 graphics card. This provided more than enough computing power for each application.

Traffic Display and Alerting

The aircraft under the control of the participant (i.e., ownship), and the surrounding aircraft were graphically depicted on a prototype Cockpit Display of Traffic Information (CDTI) developed at NASA Ames (Johnson et al., 1997). Pilots were able to scale the display from a minimum of a 10 nmi range up to 620 nmi. A multimodal alert was supported by the CDTI. A single chime would sound at the onset of an alert. Also, the aircraft involved in the predicted conflict (one of which had to be ownship), would turn from white to amber and project amber lines signifying future position indicating where the conflict was predicted to occur.

Procedure

Each participant began the session by reading and signing a form consenting to participate in the study. They were then asked to complete a short profile related to their flying experience and given a printed summary of free flight procedures. A power point presentation informing them of their task followed. Seven training runs were then administered to familiarize the pilots with the traffic display (Figure 1) and data collection interface, called the ManeuverApp (Figure 2).

For each trial, pilots were required to monitor the progress of their aircraft on the traffic display. They were responsible for making all maneuvers needed to maintain separation (specified as a lateral distance of 5nmi or vertical separation of 1000ft.), from all other aircraft.

The pilots were told that an automated alerting function was integrated into the display and they gained experience with the alert in training. They were told only that the certainty threshold of the alert would vary; the look-ahead time variation was not discussed. The look-ahead time paired with each alert was meant to appear as a manifestation of the particular conflict. The participants were not told that look-ahead time was a controlled variable.

Figure 1. Cockpit Display of Traffic Information (CDTI) in an alert state.

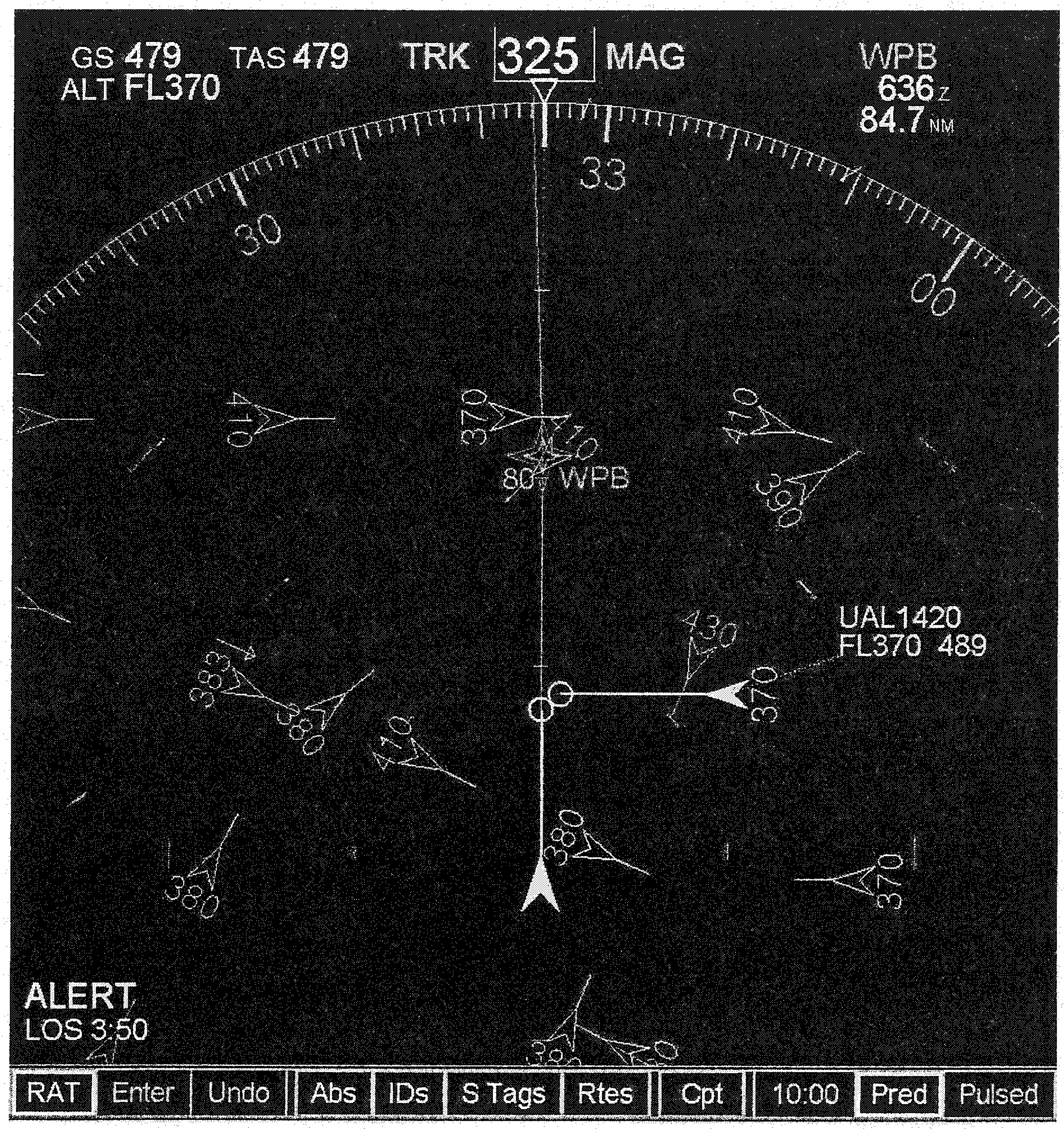
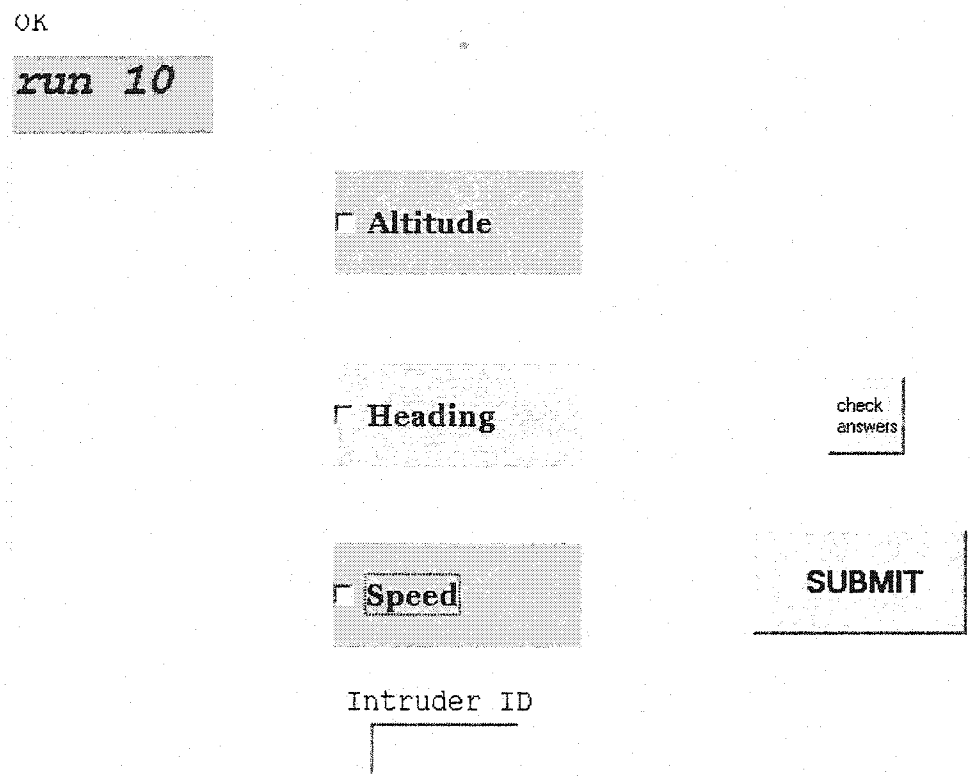


Figure 2. ManeuverApp. Used in the experiment to collect time of maneuver, type of maneuver, and intruder identification.



The pilots were given 6 trials in which the alerting certainty threshold was set at 99% and 6 trials at 75% certainty. Certainty trials were not mixed – all six certainty level runs were performed in sequence before administering the second certainty condition.

A break was offered between the sessions. As the participants monitored the scenarios, it was up to them to decide if and when a maneuver was necessary.

They were informed that they were free to maneuver at their discretion. They did not have to wait for an alert to maneuver, nor were they required to maneuver when an alert was issued. If they did decide to maneuver, they simply entered their choice of altitude, heading or speed and the scenario was quickly terminated. For each trial, the time at which the pilot initiated a maneuver was recorded along with the intruder ID.

After completing all twelve data collection runs, the pilots were asked to complete a short questionnaire (Appendix A) and an informal debrief followed.

The length of each scenario varied depending on the decisions and actions of the participants as well as the fixed script for each trial.

The resulting design was a 3 (look-ahead time) x 2 (certainty level) within subjects factorial. All twelve participants were exposed to two runs of each look-ahead/certainty combination for a total of twelve runs.

Independent Variables

Look-ahead Time: The amount of time available between the onset of an alert and the occurrence of a critical event (i.e., loss of separation [LOS]). Three levels of look-ahead time were used in this experiment: 2, 4, and 8 minutes.

Certainty: The confidence of the prediction for a given conflict alert. Two levels of certainty were used: high = 99%, low =75%.

Dependent Measures

All dependent measures were based on the time a pilot initiated the avoidance maneuver. Three different dependent measures were calculated based on different time references.

0Time: This time was measured from the start of the scenario (t=0) until the scenario was terminated. 0Time measured the total time of each scenario.

AlertTime: This was the time the maneuver was made relative to the time the alert was issued. A maneuver made before an alert was issued resulted in a negative value for AlertTime.

LOSTime: LOSTime was the amount of time prior to a loss of separation event at which a maneuver was made. A loss of separation refers to the critical event in the scenario (at which point the aircraft have breeched the prescribed minimum safety distance). When an alert was issued, it counted down the time until the loss of separation occurred.

Pilot Preference: A subjective measure of pilot preference was collected in the debrief. Pilots were asked how much look-ahead time they preferred to have available and if they believed more time would be necessary under higher workload conditions.

3. RESULTS

The first hypothesis suggested that at the late stages of conflict (2-minute condition), when there is less time for a resolution maneuver, look-ahead time would supercede alert certainty. Should this hypothesis prove true, the timing of the alert would produce significant effects ($\alpha = 0.05$). This hypothesis was supported. A significant effect for look-ahead time was observed for each variable, F_{0Time} (2,22) = 12.10, p<0.001; $F_{AlertTime}$ (2,22) = 14.19, p<0.001; $F_{LOSTime}$ (2,22) = 177.75, p<0.001. The means for all three dependent measures are summarized in Table 1. The ANOVA results are presented in Appendix B.

Table 1. Mean pilot maneuver times (and standard deviations) for each dependent measure relative to 1) start of scenario - *0Time* 2) the alert -*AlertTime* 3) the loss of separation event – *LOSTime.*

	2min(s) mean (SD)	4min(s) mean (SD)	8min(s) mean (SD)
0Time	100s (31s)	125s (39s)	174s (74s)
AlertTime	-19s (31s)	0.5s (39s)	50s (71s)
LOSTime	142s (31s)	242s (39s)	427s (73s)

The second hypothesis suggested that in the middle stages of conflict resolution (4-minute condition) certainty would play a role in the decision to maneuver. Support for this would result in differences between the certainty groups. But this hypothesis was unsupported by the experiment as certainty did not produce a significant effect $F_{0Time}(1,11) = 1.27$, $p=0.28$; $F_{AlertTime}(1,11) = 2.05$, $p=0.18$; $F_{LOSTime}(1,11) = 1.19$, $p=0.29$. These data can be seen in Appendix B

The third hypothesis, which suggested look-ahead time was more important than certainty in the early stages of conflict (8 minutes), was upheld by the same results that supported the first hypothesis. Look-ahead time produced significant effects, while the certainty levels did not produce observable significant differences. The reader is again referred to Table 1 and Appendix B. Appendix C contains graphical representations of the means for each dependent measure. Figure 3 presents the maneuver times by look-ahead time graphically.

Finally, subjective data were collected in the form of a post-run questionnaire (Appendix A). A summary of the subjective data is presented in Table 2. Pilots were asked how much look-ahead time they preferred to have for conflict resolution. The average of the 12 responses was 8.5 minutes (SD=1.3).

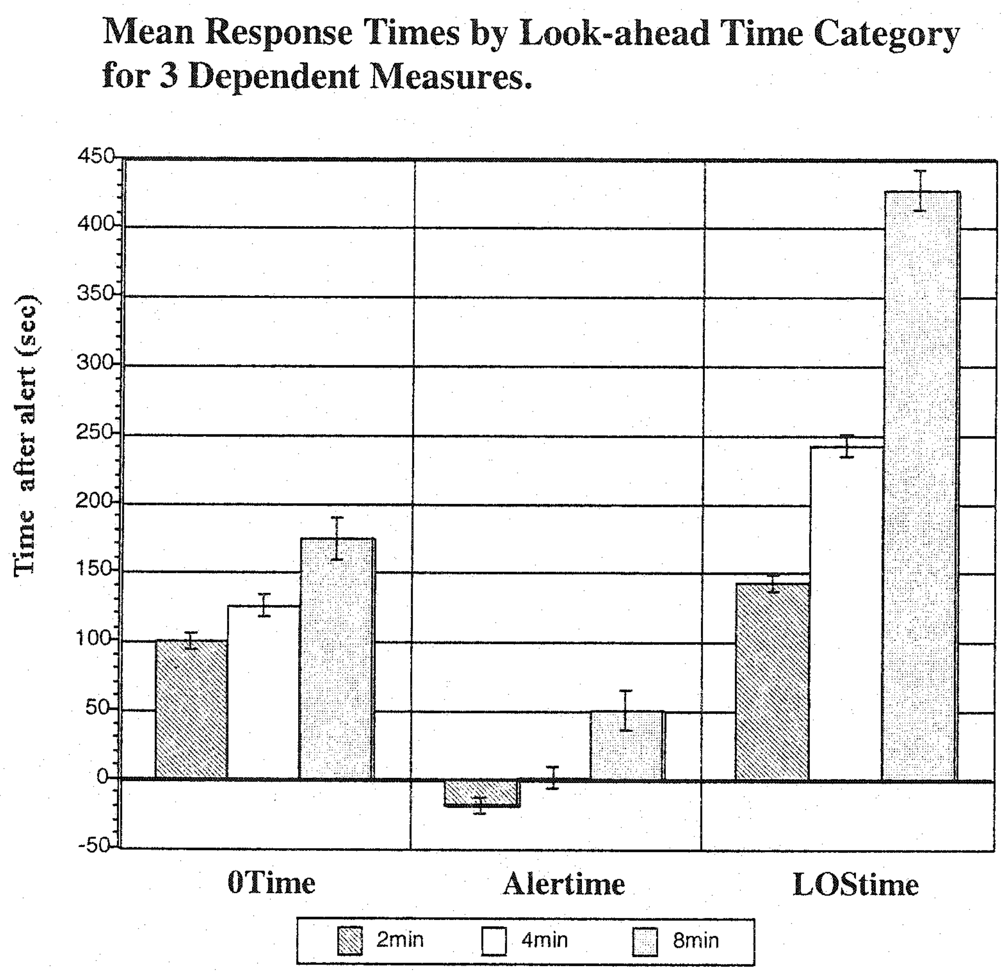


Figure 3. Means and standard error bars for maneuver times shown for each dependent variable (0Time, AlertTime, and LOSTime) by each look-ahead time level.

Table 2. Subjective responses for each participant from post-run questionnaire.

Preferred look-ahead time (min)	More time needed under higher workload	Did certainty level play a role in decisions
12	N	Y
8	N	Y
5	Y	N
7	N	N
5	N	Y
7	N	N
10	Y	Y
5	N	Y
20	Y	N
10	Y	Y
5	N	N
5	Y	Y

Mean preferred look-ahead time = 8.5 minutes (SD=1.3min)

4. DISCUSSION

There were three important results found in this study. The first result demonstrated that on average the pilots in this study maneuvered *before* the alert was issued in the late stages of conflict resolution (2-minute condition). Second, the different certainty levels (75% and 99%) did not seem to affect the decisions made by the pilots. Finally, an extended look-ahead time (8-minute condition), did show a significant delay in reacting to the alert. These results however, do not produce the interaction of look-ahead time and alert certainty as predicted by the hypotheses.

To investigate the late stages of conflict resolution, a 2-minute look-ahead time was selected. Based on the work of Jürgensohn et al. (2001), it was hypothesized that in the late stages of conflict, pilot decisions would be governed more by the look-ahead time than by the certainty of conflict. This hypothesis was supported by the data as the certainty levels did not produce a significant difference while the look-ahead time did impact behavior significantly. The average of the response times in the 2-minute condition showed that on average, pilots initiated maneuvers 19 seconds (SD=31) before the onset of the alert. These results are comparable to the "late alert" data obtained by Jürgensohn et al. (2001) in the late stages of their controller task.

For look-ahead times of two minutes or less, the conflicting aircraft were separated by less than 20 nmi. This was easily recognized on the traffic display as the conflict aircraft would be well inside the 20 nmi range ring (range rings are concentric circles showing relative distance from ownship). Perhaps the proximity of the two aircraft and their crossing trajectories was sufficient to instigate the decision to maneuver without receiving a conflict alert.

In general, the participants were to believe the low level alert threshold would provide more advanced notice of a conflict, and be less likely to "miss" (fail to identify) a conflict. If the pilots received and trusted this alert, they would be expected to maneuver later in the low certainty condition than they might in the high certainty condition (where the possibility of a missed alert was greater). But a certainty difference failed to emerge in the pilots' behavior. At both certainty levels, pilots maneuvered before the alert in the 2-minute condition. These results indicate two minutes would represent a late alert to the pilots in this study. Further, as the trials in this study were terminated immediately following a maneuver, no alert was received. Thus, the 2-minute alert condition was seen as a missed alert from the perspective of most of the pilots.

The second finding in this study indicated that certainty level had no observable influence on pilot decision making. It was believed a difference could result in the middle stage of conflict resolution (4 minutes), but it was not obtained. Pilots reacted similarly to each level of certainty in all look-ahead time conditions. In the post-run questionnaire, five of the twelve pilots indicated the certainty level did not play a role in their decisions (Table 2). Gempler and Wickens (1998) similarly failed to observe a main effect of certainty.

During the debrief, all pilots suggested their decision making was conservative; they preferred to move if there was any possibility of conflict. They suggested safety was paramount and chose to maneuver even if it might not have been completely necessary.

Issues such as wake turbulence and passenger comfort took precedent over flight efficiency. Several pilots commented they would not feel comfortable flying near the minimum separation distance prescribed in the study (5nmi laterally, 1000 ft. vertically). This feedback suggests that conservative decision makers in this study viewed 75% certainty as sufficient reason for maneuver initiation. The cost of an unnecessary maneuver was far less than the cost of failing to maneuver when a course deviaiton was required.

The third result demonstrated a response latency in the early stages of conflict (8 minutes/ 70nmi before a critical event) as predicted. On average, maneuvers were not made until 50 (SD=71) seconds after the onset of the alert. Pilots were not compelled to maneuver immediately and instead monitored the situation for almost another minute.

Jürgensohn et al. (2001) would likely describe such an alert at this early stage to be "attention directing." This type of alert encourages, "the operator to monitor objects for which a potential conflict was indicated" (p. 459). This occurs when an operator is warned of a potential conflict "very early" (p. 459).

This is contrasted with the other type of warning discussed by Jürgensohn et al. (2001) called "response eliciting". They define such an alert as requiring intervention very soon after receiving the alert. The maneuver times collected during the 4-minute trials fit this description of response eliciting quite well. The average response was 0.5 seconds (SD=39) after the alert triggered. This is only an average and does not suggest that 4 minutes is the optimal look-ahead time for initiating pilot action. But it does seem to show a willingness on the part of the pilots to take action within this time frame. It should also be noted that selecting a maneuver requires monitoring and decision making before any goal-directed intervention can occur. Thus, cognitive processes such as attention, trajectory estimation, and decision making must have been activated before receiving an alert. The pilots decided to perform these tasks on their own without cueing from the alert.

To summarize, the look-ahead time was the only independent variable that impacted pilot decision making behavior in this free flight task. The levels of the look-ahead time produced some interesting results. The alerts that provided only a 2-minute look-ahead were representative of a system that provided late (or missed) alerts. Look-ahead times in the 4-minute range are more consistent with a response eliciting system. Finally, the 8-minute look-ahead can be likened to an attention directing alert system. Should these data be replicated, the information could be useful in designing an alerting system when specific characteristics are desired.

5. CONCLUSION

The intent of this study was to identify a time envelope that pilots preferred when confronted with conflicts at a given level of certainty. It was not intended to identify performance minima, such as the least amount of time a pilot may be capable of solving a conflict. Rather, it was hoped to identify a preference or a comfort zone in which pilots might choose to initiate conflict resolution actions.

Further research needs to be conducted in this area in order to move toward optimal alerting strategies. First, it will be necessary to decide what the function of the alert should be. Should these types of conflict alerts be attention directing or response eliciting? Should the alert be multi-staged in an effort to provide both? If so, when should the stage-one alert be issued and what should the time interval be between the stages? Where does this leave certainty? It's premature to disregard the importance certainty may have in conflict alerting. Perhaps higher fidelity simulations may discover a certainty effect.

Future studies will need to investigate these issues. The robustness of their results will be enhanced by using greater fidelity simulations than the part-task simulator used for this study. Pilots should be tasked with comparable workloads and activities they would experience while in flight. The present study afforded pilots the opportunity to monitor the traffic to a greater extent than they would likely be able in the real world. This could lead to significant performance differences between the flight deck and the lab environment.

The current study hoped to provide a baseline for pilot response under varied alerting parameters of look-ahead time and certainty. Since there was no coordination possible and the workload was minimal, the results of this experiment should be considered to be only one component of a response period required for airborne conflict management. Ideally, this element will be a building block upon which other factors can be added to gain a more accurate model of pilot response under various circumstances.

REFERENCES

Adler, E.L., & Ruelos, P.E. (1993). GPS – Global Positioning System. Using the global positioning system for navigation on Boeing airplanes. Airliner, Oct-Dec., 11-18.

Banbury, S., Selcon, S., Endsley, M., Gorton, T., & Tatlock, K. (1998). Being certain about uncertainty: how the representation of system reliability affects pilot decision making. In Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting (pp. 36-39). Santa Monica, CA: Human Factors and Ergonomics Society.

Bliss, J.P., Gilson, R.D. & Deaton, J.E. (1995). Human probability matching behaviour in response to alarms of varying reliability. Ergonomics. 38(11), 2300-2312.

Box, G.P., Hunter, W.G., Hunter, J.S. (1978). Statistics for experiments. An introduction to design, data analysis, and model building. New York: John Wiley & Sons, Inc.

Cole, R.E., Green, S., Jardin, M., Schwartz, B.E. & Benjamin, S.G. (2000). Wind prediction accuracy for air traffic management decision support tools. Proceedings of the 3rd USA/Europe Air Traffic Management R&D Seminar. Napoli, Italia.

DAG-TM Team, Advance Air Transportation Technologies Program (1999). Concept definition for distributed air/ground traffic management, version 1.0. NASA Ames Research Center, Moffett Field. [On-line] Available: http://human-factors.arc.nasa.gov/ihh/cdti/DAG_TM_WEB/dag2001.html

DiMeo, K., Kopardekar, P., Ashford, R., Lozito, S., & Mackintosh, M. (2001). Shared-separation: Empirical results and theoretical implications. Proceedings of the 4th USA/Europe Air Traffic Management R&D Seminar. Santa Fe, NM. December 2000.

Dror, I.E.,, Busemeyer, J.R., Basola, B. (1999). Decision making under time pressure: An independent test of sequential sampling methods. Memory & Cognition. 27(4), 713-725.

Gempler, K.S., & Wickens, C.D. (1998). Display of predictor relilability on a cockpit display of traffic information. Final Technical Report ARL-98-6/Rockwell-98-1

Hoekstra, J., van Gent, R., & Ruigrok, R., (1998). Conceptual design of free flight with airborne separation assurance. NLR-TP-98252. AIAA Guidance, Naviation and Control Conference. Boston, MA. August, 1998.

Hoffman, E., Nicolaon, J., Pusch, C., & Zeghal, K. (1999). Limited delegation of separation assurance to aircraft. Presented at Air Traffic Management '99 Workshop. Capri, Italy. [On-line]. Available: http://www.eurocontrol.fr/projects/freer/publications.htm

Johnson, W., Battiste, V., & Holland, S., (1999). A cockpit display designed to enable limited flight deck separation responsibility. NASA TP-1998-208439. Proceedings of the 1999 World Aviation Congress, Anaheim, CA.

Johnson, W. W., Battiste, V., Delzell, S., Holland, S., Belcher, S., & Jordan, K. (1997).Development and demonstration of a prototype free flight cockpit display of traffic information. Proceedings of the 1997 SAE/AIAA World Aviation Congress. American Institute of Aeronautics and Astronautics.

Jürgensohn, T., Ilyong, P, Sheridan, T.B., & Meyer, J. (2001). Conflict warnings and the search and detection of collisions in a simulated atc-like environment. . In Proceedings of the Human Factors and Ergonomics Society 45rd Annual Meeting (pp 458-462)., Santa Monica, CA: Human Factors and Ergonomic Society.

Kantowitz, H., Hanowski, R. J., & Kantowitz, S.C. (1997). Driver acceptance of unreliable traffic information in familiar and unfamiliar settings. Human Factors 39(2), 164-176.

Kremer, H., Bakker, B., & Blom, H. (1997). Probabalistic versus geometric conflict probing. [On-line] Available : www.atm-seminar.97.eurocontontrol.fr/kremmer.htm.

Kuchar, J. (2001). Managing uncertainty in decision-aiding and alerting system design. Proceedings of the 6th CNS/ATM Conference, Taipei, Taiwan.

Kuchar, J., & Hansman, R.J., Jr., (1995). A unified methodology for the evaluation of hazard alerting systems (Report No. ASL-95-1). Cambridge, MA. Massachusetts Institute of Technology, International Center for Air Transportation.

Lourerio, T., Blin, K., Hoffman, E., and Zeghal, K. (2001). Development of a tool for comparing conflict detection algorithms for air traffic management (AIAA-2001-4053). [On-line] Available: www.eurocontrol.fr/projects/freer/archive/GNC-2001-4053.pdf.

Mackintosh, M., Dunbar, M., Lozito, S., Cashion, P., McGann, A., Dulchinos, V., Ruigrok, R., Hoekstra, J., & van Gent, R. (1998). Self-separation from the air and ground perspective. Proceedings of the 2nd USA/Europe Air Traffic Management R&D Seminar. Orlando, FL. 1998.

Maule, A.J., Edland, A.C. (1997). Time pressure and judgement. In R. Raynard, W.R. Crozier & O. Svenson (Eds.), Decision making: Cognitive models and explanation (pp. 189-204). New York, NY: Routledge

Metzger, U., & Parasuraman, R. (1999). Free Flight and the air traffic controller: Active control versus passive monitoring. In Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting . Santa Monica, CA: Human Factors and Ergonomics Society.

Paielli, R., & Erzberger, H. (1997). Conflict probability estimation for free flight. Journal of Guidance, Control, and Dynamics. Vol. 20, No. 3, May-June 1997.

Parasuraman, R. (1997). Humans and automation: Use, misuse, disuse, abuse. Human Factors, 39(2), 230-253.

RTCA. (1995). Final report of RTCA Task Force 3 free flight implementation. Washington, D.C.: RTCA

RTCA Special Committee 186 (2000). Application of Airborne Conflict Management: Detection, Prevention, & Resolution (DO-263) Prepared by SC-186. Washington, D.C.: RTCA.

RTCA Special Committee 186, WG-1, Airborne Conflict Management Sub-Group. (2002). Airborne conflict management. Application description. Version 2.5. March 2002. Unpublished manuscript.

Smith, K., Briggs, A., Knect, W., & Hancock, P. (1997). Success and failure at self-separation in simulated free flight. Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting, Santa Monica, CA. Human Factors and Ergonomics Society. pp. 13-17.

Stenbit, J. P. (2001). Global Positioning System Standard Positioning Service Performance Standard. OASD(c31)/ODASD(C31SR&S)/Space Systems. [On-line] Available: http://www.navcen.uscg.gov/gps/geninfo/2001SPSPerformanceStandardFINAL.pdf

Williams, D.H. & Green, S.M. (1998). Flight Evaluation of Center-Tracon Automation System Trajectory Prediction Process (NASA Report NASA/ TP 1998-208439)

Yang, L.C. & Kuchar, J.K. (1997). Prototype conflict alerting system for free flight. Journal of Guidance and Control. Vol. 20, No. 4. July-August 1997.

APPENDIX A

Post-Run Questionnaire

Post Run Questionnaire
Circle one of the two answers, or write in your answer on the line provided.
Comments may be added to the back if more room is needed.

1) Do you feel the certainty level influenced your decisions? **Y** **N**
Please elaborate:

2) Assume this technology was actually to be implemented on your aircraft in everyday flying conditions.

Which certainty threshold would you prefer? **High** **Low**

Which alerting imprecision is more tolerable? **Missed alert** **False alert**
Please elaborate:

3) In this experiment, the workload is considered to be extremely low.
Under these conditions, how much look-ahead time would you like the alerting logic to provide? (in minutes) ___________min.

Do you think more time would be needed under higher workload conditions? **Y** **N**
Please elaborate:

4) The scenarios were ended before you ever had the chance to discover if a conflict actually existed (a true loss of separation would have occurred).

What percentage of the trials do you believe contained a true conflict? (any number between 0% - 100%). ___________
Please elaborate:

APPENDIX B

ANOVA Tables

Table 3. ANOVA output for dependent measure 0Time.

SOURCE: grand mean

time	cert	N	MEAN	SD	SE
		72	133.6250	59.6811	7.0335

SOURCE: time

time	cert	N	MEAN	SD	SE
2		24	100.9792	31.9730	6.5265
4		24	125.0833	39.0595	7.9730
8		24	174.8125	74.0873	15.1230

SOURCE: cert

time	cert	N	MEAN	SD	SE
	99	36	130.2917	61.0003	10.1667
	75	36	136.9583	59.0046	9.8341

SOURCE: time cert

time	cert	N	MEAN	SD	SE
2	99	12	100.9167	30.1224	8.6956
2	75	12	101.0417	35.0729	10.1247
4	99	12	121.7500	41.2016	11.8939
4	75	12	128.4167	38.3168	11.0611
8	99	12	168.2083	81.5195	23.5327
8	75	12	181.4167	68.8205	19.8668

FACTOR:	part	time	cert	0time
LEVELS:	12	3	2	72
TYPE :	RANDOM	WITHIN	WITHIN	DATA

SOURCE	SS	df	MS	F	p
mean	1285606.12	1	1285606.12	135.25	0.000
p/	104552.54	11	9504.77		
time	**68042.89**	**2**	**34021.44**	**12.10**	**0.000**
tp/	**61852.93**	**22**	**2811.49**		
cert	800.00	1	800.00	1.27	0.283
cp/	6918.50	11	628.95		
tc	513.52	2	256.76	0.55	0.583
tcp/	10209.97	22	464.09		

Table 4. ANOVA output for dependent measure AlertTime

SOURCE: grand mean

time	cert	N	MEAN	SD	SE
		72	10.5278	58.0254	6.8384

SOURCE: time

time	cert	N	MEAN	SD	SE
2		24	-19.7708	31.5839	6.4470
4		24	0.5417	39.5914	8.0816
8		24	50.8125	71.2263	14.5390

SOURCE: cert

time	cert	N	MEAN	SD	SE
	99	36	5.6111	56.5208	9.4201
	75	36	15.4444	59.8812	9.9802

SOURCE: time cert

time	cert	N	MEAN	SD	SE
2	99	12	-19.8333	30.6233	8.8402
2	75	12	-19.7083	33.8818	9.7808
4	99	12	-2.9167	41.3565	11.9386
4	75	12	4.0000	39.2556	11.3321
8	99	12	39.5833	73.9720	21.3539
8	75	12	62.0417	69.7181	20.1259

FACTOR:	part	time	cert	alert
LEVELS:	12	3	2	72
TYPE :	RANDOM	WITHIN	WITHIN	DATA

SOURCE	SS	df	MS	F	p	
mean	7980.05	1	7980.05	0.87	0.370	
p/	100582.86	11	9143.89			
time	**63374.09**	**2**	**31687.04**	**14.19**	**0.000**	***
tp/	**49125.99**	**22**	**2232.99**			
cert	1740.50	1	1740.50	2.05	0.179	
cp/	9298.08	11	845.28			
tc	1572.89	2	786.44	1.29	0.294	
tcp/	13358.52	22	607.20			

Table 5. ANOVA output for dependent measure LOSTime

SOURCE: grand mean

time	cert	N	MEAN	SD	SE
		72	270.9792	129.0947	15.2140

SOURCE: time

time	cert	N	MEAN	SD	SE
2		24	142.7708	31.5839	6.4470
4		24	242.9167	39.3274	8.0277
8		24	427.2500	73.7913	15.0626

SOURCE: cert

time	cert	N	MEAN	SD	SE
	99	36	274.2083	133.2655	22.2109
	75	36	267.7500	126.5933	21.0989

SOURCE: time cert

time	cert	N	MEAN	SD	SE
2	99	12	142.8333	30.6233	8.8402
2	75	12	142.7083	33.8818	9.7808
4	99	12	246.2500	39.5144	11.4068
4	75	12	239.5833	40.5988	11.7199
8	99	12	433.5417	81.4420	23.5103
8	75	12	420.9583	68.3093	19.7192

FACTOR:	part	time	cert	loss
LEVELS:	12	3	2	72
TYPE :	RANDOM	WITHIN	WITHIN	DATA

SOURCE	SS	df	MS	F	p	
mean	5286939.03	1	5286939.03	559.46	0.000	***
p/	103949.26	11	9449.93			
time	**999490.89**	**2**	**499745.44**	**177.75**	**0.000**	*******
tp/	**61852.93**	**22**	**2811.49**			
cert	750.78	1	750.78	1.19	0.29	
cp/	6897.67	11	627.06			
tc	466.02	2	233.01	0.52	0.60	
tcp/	9838.14	22	447.18			

APPENDIX C

Graphical Representation of the Means for Each Dependent Measure

Figure 4. Means and standard error bars for elapsed times from scenario start at which pilots initiated resolution maneuvers for each look-ahead time level.

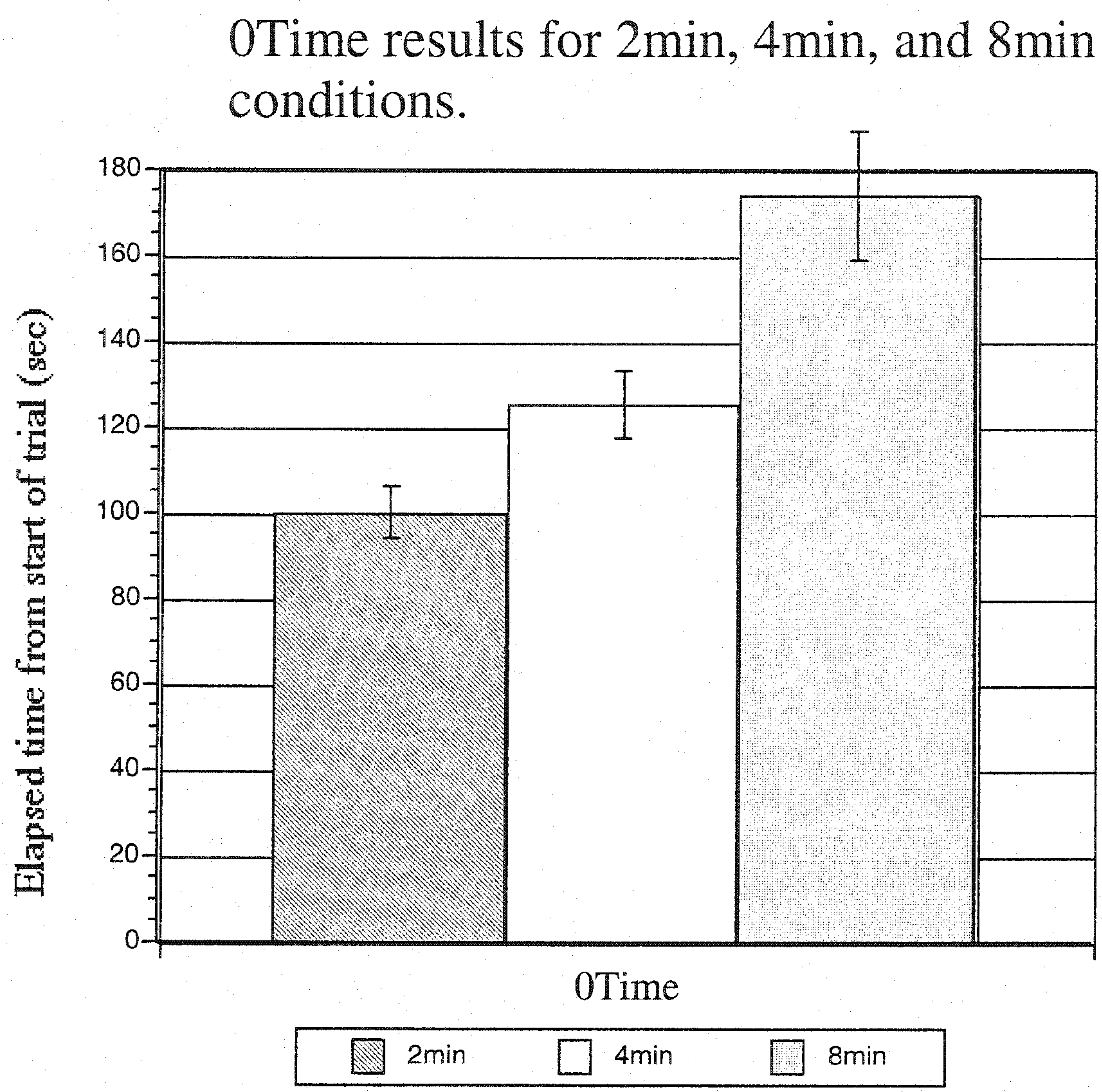


Figure 5. Means and standard error bars for maneuver times relative to the alert for each look-ahead time level (a negative value indicates the maneuver occurred before an alert was issued).

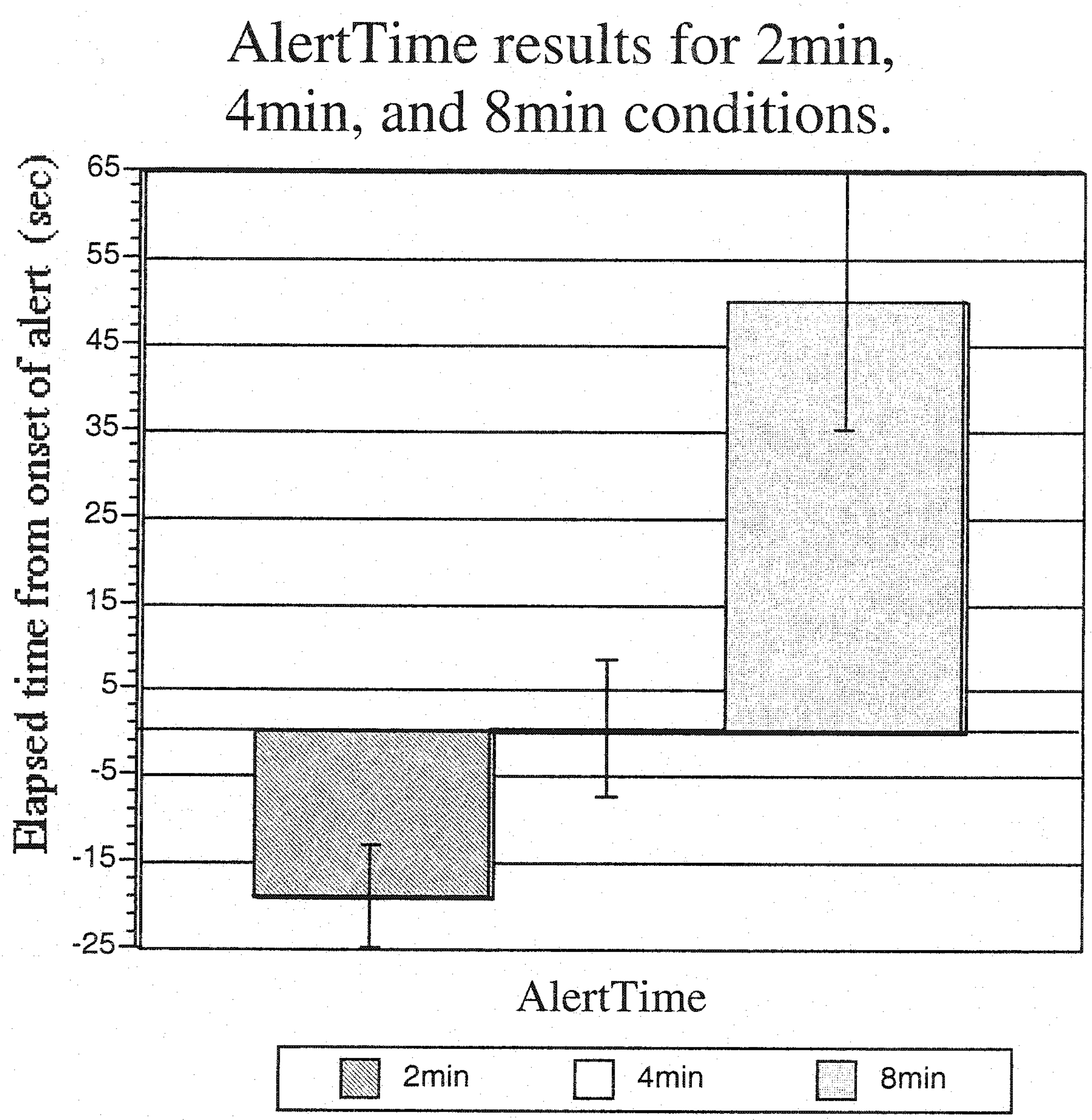


Figure 6. Means and standard error bars for number of seconds prior to the loss of separation event the maneuver was initiated for each look-ahead time level.

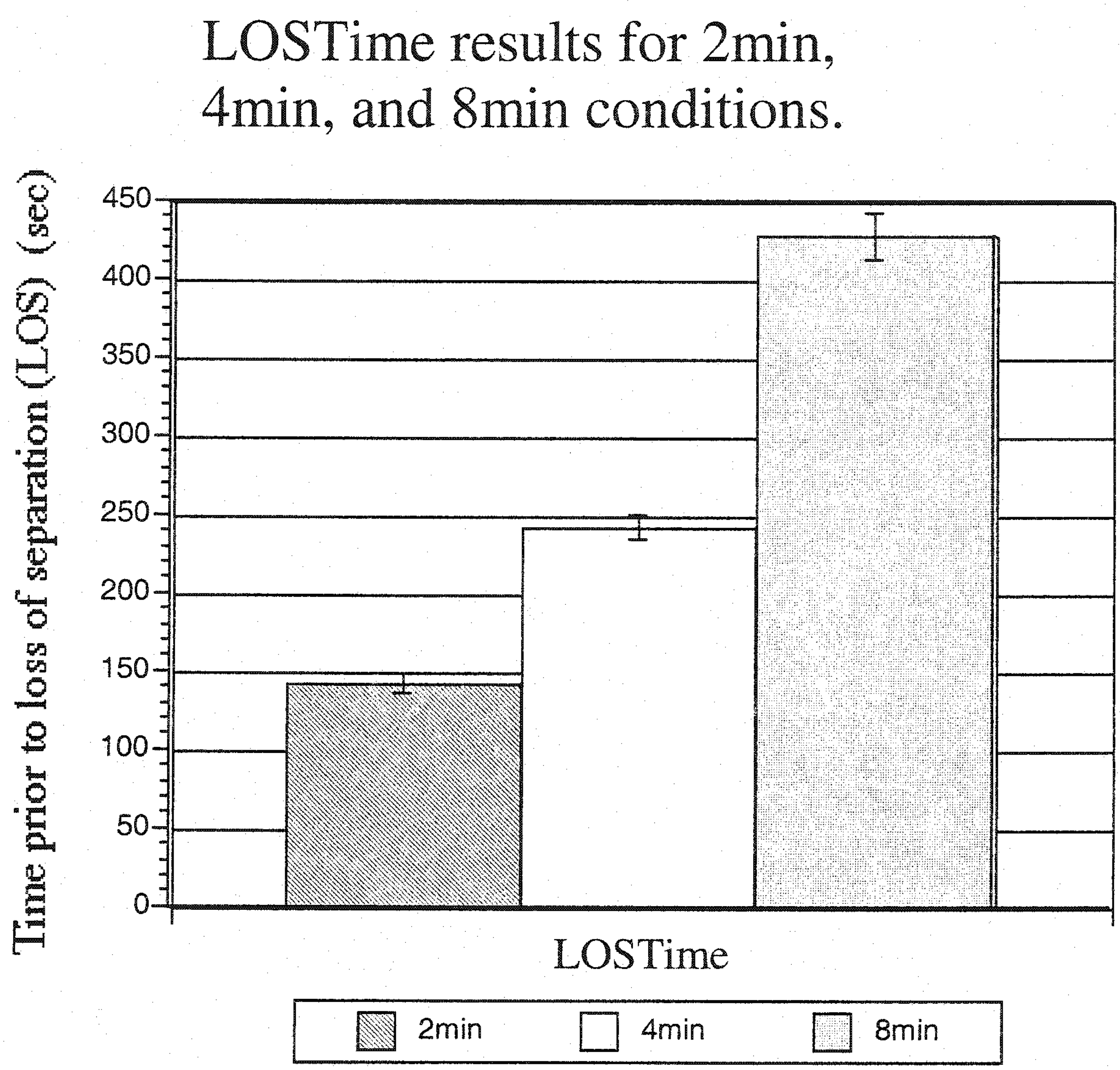


Figure 7. Means and standard error bars for elapsed time from scenario start at which pilots initiated resolution maneuvers for each level of alert certainty.

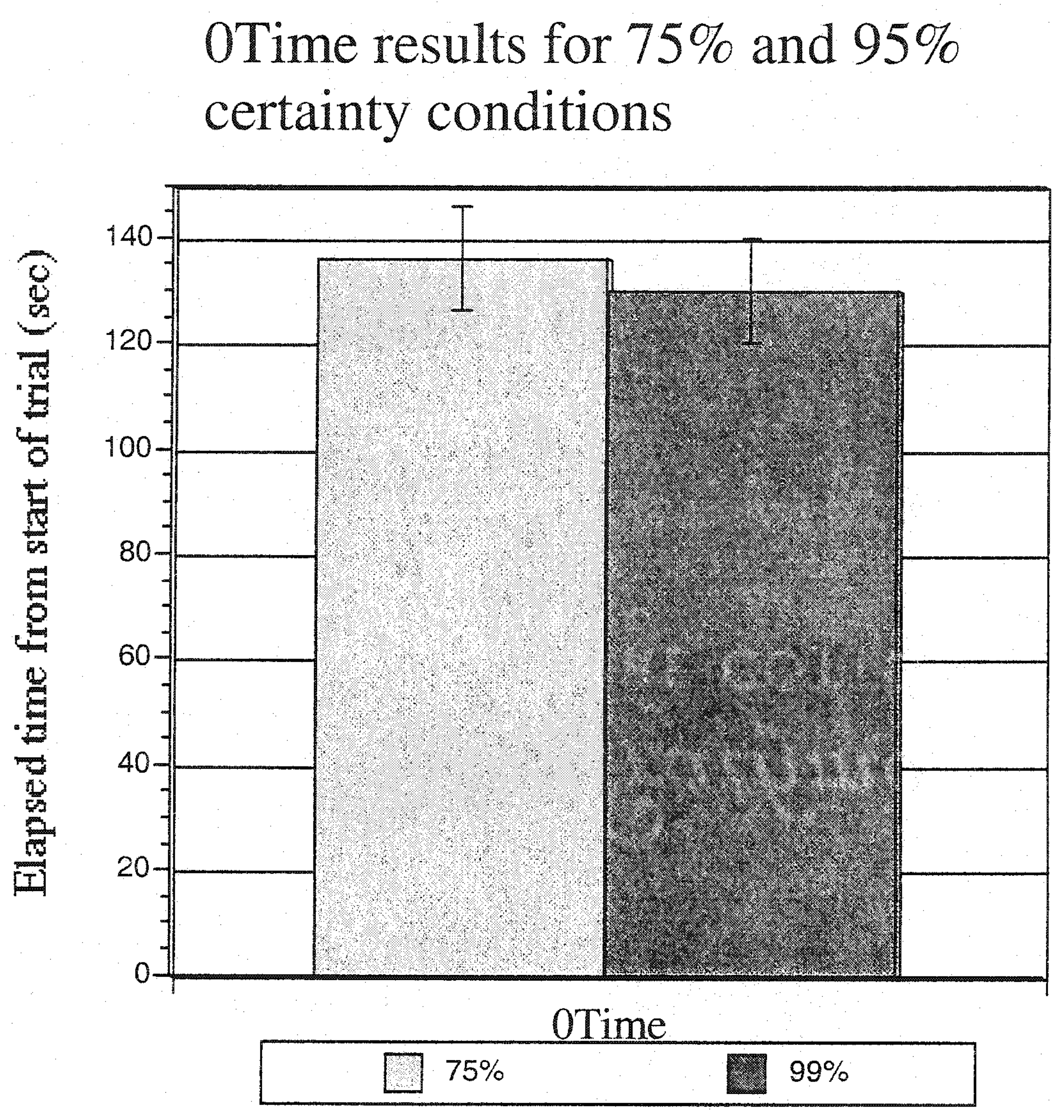


Figure 8. Means and standard error bars for maneuver time relative to the alert for each certainty level.

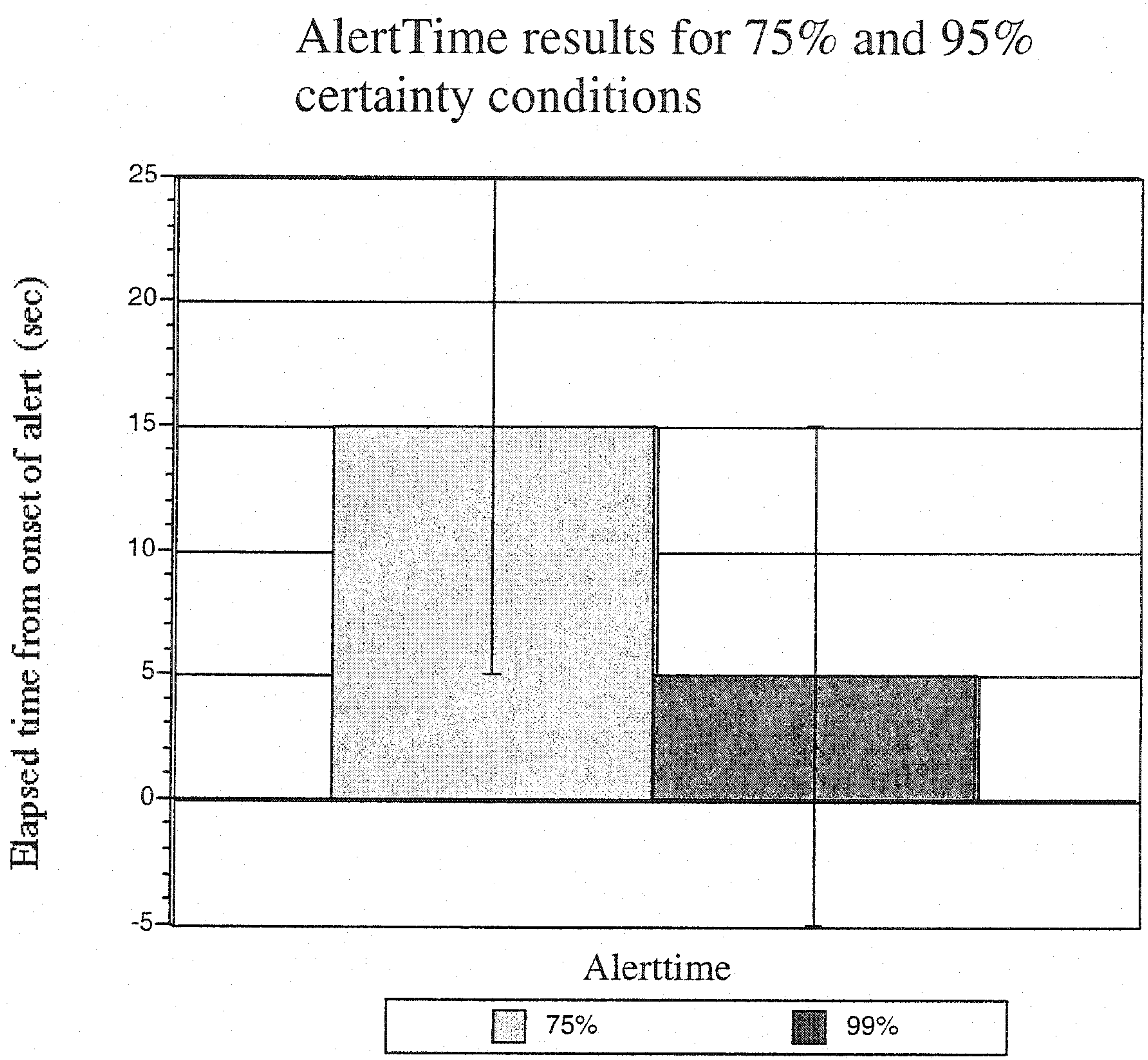
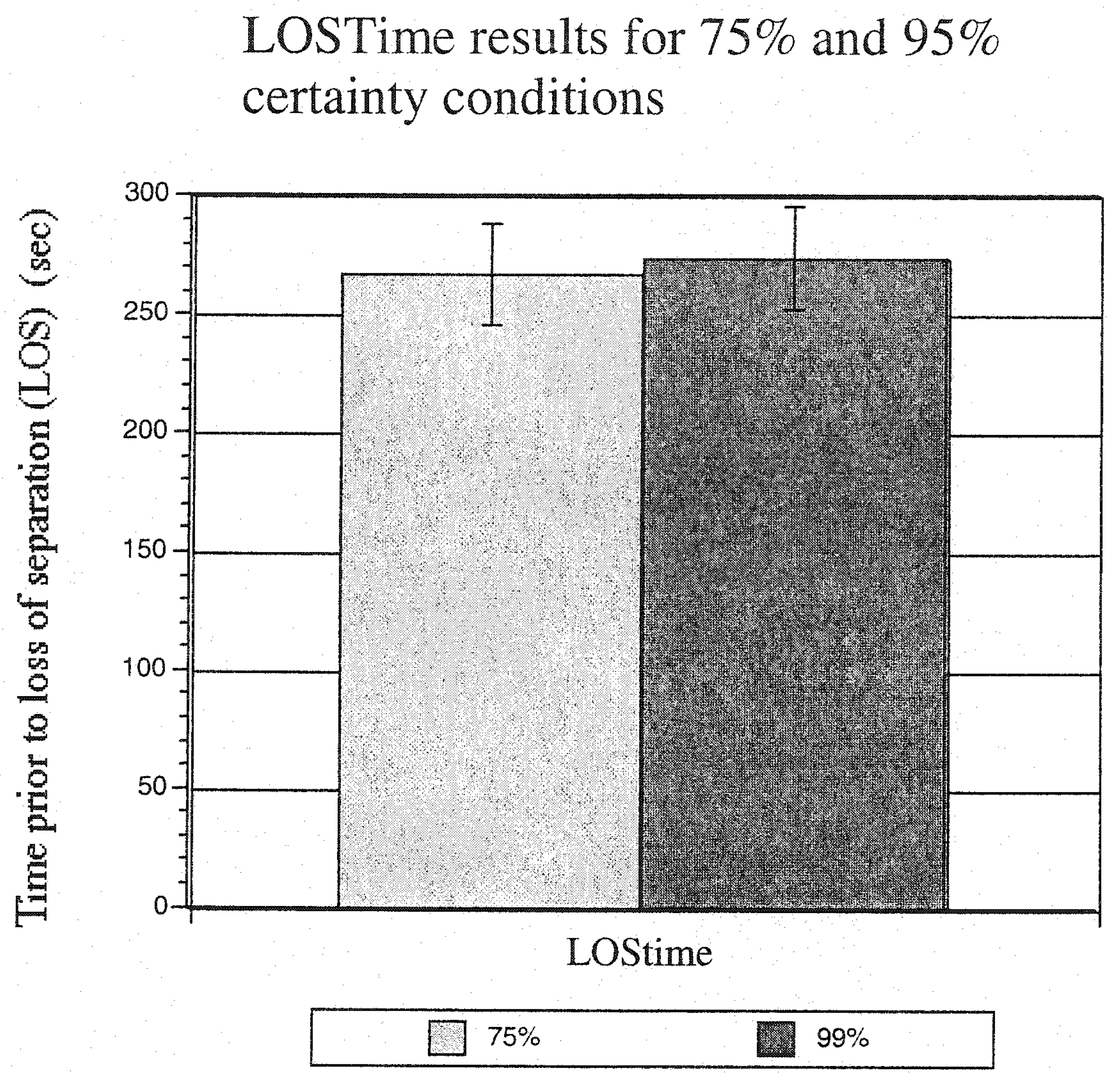


Figure 9. Means and standard error bars for number of seconds prior to the loss of separation event the maneuver was initiated for each certainty level.



APPENDIX D

Conflicts and Alerting Algorithms

Conflicts.

Two aircraft are said to be in geometric conflict when the distance between them (which is the distance measured between their protected zones) has been reduced to less than a prescribed minimum distance (Kremer, Bakker, & Blom 1997; Loureiro, Blin, Hoffman, & Zeghal 2001).

As noted by Loureiro et al. (2001), an algorithm dependent only on geometric prediction fails to account for the connate uncertainty in such complex calculations. With the geometric method, trajectories are predicted on current state (velocity) vector and may possess intent information (such as from a flight plan). The predictions are extrapolated forward and a determination is made whether any points along the two trajectories are separated by less than a prescribed minimum. Should it be predicted that the minimum threshold will be violated, an alert is issued. It is acknowledged that such an approach can be sufficient over short distances and minimal time frames. However, as this method does not account for the uncertainties discussed, longer time frames are likely to produce greater error in prediction. A further consequence of the less robust geometric approach was its inferior performance; it resulted in more false alarms and more missed alerts compared to other methods (Loureiro et al., 2001; Kremer et al., 1997; Kuchar & Hansman 1995).

In the current experiment a probabilistic algorithm was used to evaluate traffic threats. The algorithm, developed by Yang and Kuchar (1997) at the Massachusetts Institute of Technology, utilized a probabilistic approach (as opposed to geometric) to assess the existence of a conflict. This method fosters an ability to cope with uncertainty in conflict prediction. A further advantage of a probabilistic approach is that it affords the opportunity to assess the system in terms of safety and false alarm probabilities (Kuchar 2001).

<u>Alerting Algorithms</u>

Paielli and Erzberger (1997) devised an analytical means of predicting a loss of separation should aircraft continue along conflicting trajectories. But, as it has already been stated, these predictions contain error. For each aircraft trajectory, along-track and cross-track errors can be calculated.

The cross-track error is generally the more stable of the two. Aircraft with automated navigation systems typically do not exceed 0.5 nmi variations in their cross-track position over a 30 minute period. Aircraft lacking such sophistication are more unstable but the skill of the pilot can usually constrain the error to 1 nmi or less (Paielli & Erzberger 1997).

It is the along-track error which generates the most uncertainty. Paielli and Erzberger (1997) suggested the growth rate was linear and typically a 0.25 nmi/min incongruity in the along-track direction. This is largely due to imprecise wind information. Further degrading the along-track certainty is the capacity for the error to accumulate in this dimension as the flight progresses. While cross-track error may cancel itself out (either by autopilot or manual control as left/right deviations serve to nullify each other), along-track perturbations are often additive.

In order to deal with the uncertainty, Paielli and Erzberger (1997) approximate all errors as normally distributed (Gaussian). Using the normally distributed error assumption, "error ellipses" can be constructed for each aircraft. The principal axis of the ellipse lies in the along-track direction. This highly analytical method of prediction is limited in its ability to account for non-Gaussian error.

To address complexities beyond a Gaussian idealization, Yang & Kuchar (1997) employed a simulation technique (Monte Carlo) as the basis for their alerting logic. Consequently, the processing time and computing requirements significantly increased. While linear conflict scenarios yield equivalent results for both alerting logics, the Yang-Kuchar logic provided a more robust algorithm capable of handling a broader conflict set

For a given encounter, thousands of simulations are run to establish a baseline probability of a loss of separation. The escape maneuvers available are also assessed. A combination of the likelihood of conflict and the avoidance maneuvers available to the crew then results in a particular level of alert (only the highest alert level was used for the present study).

A further enhancement of the Yang-Kuchar logic was that it was developed specifically for airborne use where conflicts are solved on the flight deck. A protected zone (based on the separation standard) was constructed for each aircraft. In their case a cylinder with a 5nmi diameter and 2000ft height ($\pm$ 1000ft above and below the aircraft) was used.

A point of interest related directly to this study is a distribution used by Yang and Kuchar (1997) to approximate pilot latency in reacting to an alert. They employed a gamma distribution to represent pilot reaction times. They suggested 95% probability that the response would be within two minutes from the onset of the alert. Two minutes was selected to permit coordination time with other airborne or ground participants.

While this two minute time frame seems logical enough, there is little empirical data to support it as this characteristic has rarely been the focus of investigation. Additionally, the latency in response is likely to fluctuate with many variables such as workload, traffic density, severity of alert, communications availability, conflict geometry, and phase of flight to name only a few.

The current study hoped to provide a baseline for pilot response under varied alerting parameters of look-ahead time and certainty. Since there is no coordination possible and the workload is minimal, the results of this experiment should be considered to be only one component of a response period. Ideally, this element will be a building block upon which other factors can be added to gain a more accurate model of pilot response under various circumstances.

APPENDIX E

Uncertainty Due to Winds and Navigation

Navigation

Even with the deployment and utilization of the Global Positioning System (GPS), instantaneous position cannot be known precisely. In fact, it was not until midnight, May 1, 2000, that the civilian population was permitted to leverage the full capabilities of the Defense Department's satellite constellation (Stenbit 2001). Prior to this directive, civilian use of GPS data was intentionally degraded (via Selective Availability (SA)) such that true horizontal position was accurate to only 100 meters 95% of the time, and to 300 meters 99.99% of the time (Adler & Ruelos 1993). The vertical accuracy was reported within 140 meters with 95% reliability.

While a substantial improvement has been garnered by the civil aviation community, GPS monitoring is still imperfect. In fact, in the GPS System Standard Positioning Service Performance Standard (Stenbit 2001), the Assistant Secretary of Defense outlines several error sources "not under direct control of GPS constellation operations" (p 8). He specifically names ionosphere and troposhere model errors, signal reception interference, and receiver thermal noise, as unmitigated saboteurs in the system.

Despite these errors, the position accuracy is quite suitable for navigation. The Global Average Positioning Domain Accuracy achieves $\leq$ 13 meters horizontally and $\leq$ 22 meters vertically with 95% reliability. It is stressed here that these numbers apply only to the instantaneous position of the aircraft. Trajectory prediction is a more toilsome task.

Wind Inaccuracies

Cole, Green, Jardin, Schwartz, and Benjamin (2000) reported on wind prediction inaccuracies based on flight tests in 1992 and 1994. "Flight tests have indicated that wind prediction errors may represent the largest source of trajectory prediction error" (page 2). Two versions of he Rapid Update Cycle (RUC) were used to calculate wind predictions. The version employed in 1994 was run every hour. Version 1, used in 1992 was run every three hours. Both systems provided hourly forecasts for up to 12 hours. The data collection equipment available for each test phase included: ACARS equipped commercial aircraft, wind profiles from vertically oriented radar, surface observations, as well as wind and moisture readings from weather satellites.

This data was then utilized by the ground-side Decision Support Tool (DST) called the Center-TRACON Automation System (CTAS). Winds aloft data were updated every three hours. CTAS converted the data to match Denver center (ZDV) coordinates. With this information, CTAS computed its own winds forecast as well as the flight path for each aircraft.

In order to compare the system's prediction capabilities, a means of gathering current, accurate data was needed. That was accomplished by NASA's Transport Systems Research Vehicle (TSRV). The TSRV is a Boeing 737 equipped with GPS and the "flight-test air data system". Over both Phase I and Phase II fifty test runs were recorded. Each run consisted of a 100 –200 nmi arrival path which started in cruise (FL330 – FL350) and followed a descent to about 17,000 feet The data collected by the

TSRV served as the true values for a given location at a specified time. The CTAS forecast data was judged relative to these values.

The discrepancies elucidated by these data clearly demonstrate the difficulty in making accurate predictions and the inherent uncertainty in conflict alerting algorithms. Errors of greater than 20 knots were detected in some areas. The 20 knot errors were typically found in cruise where there is a cumulative effect with time. One series of test runs consistently experienced a 60 knot differential (this was the result of a passing front which was not forecast correctly). A headwind calculation 15 knots in error, projected over a 20 minute flight segment, is sufficient to create a 5 nmi discrepancy.

Williams and Green (1998) identified further inaccuracies beyond the calculation of wind predictions. They suggested that in addition to employing imperfect information there is also a time lag. The addition of the time delay makes the data less powerful in modeling current conditions. A modern FMS permits the entry of wind and temperature data for each waypoint. There is also a provision to input weather at different altitudes appropriate for a descent profile. However, these data are calculated and input prior to take-off. The occurrence of in-flight updates to this information is infrequent (Williams & Green 1998). Thus, during longer flights, information that was 3-6 hours old at entry cannot be guaranteed to accurately represent the current environment.

With all of these factors degrading the fidelity of trajectory predictions, the certainty of a conflict should be considered dynamic.

APPENDIX F

Scenario Construction

The primary manipulations of this study were the look-ahead time to a conflict and the associated certainty of the annunciated alert. In air traffic management many other variables are known to have an affect on performance. These include: conflict geometry, the number of aircraft in a scenario, right of way rules for conflict resolution, traffic flow, and airspace configuration. In order to maintain experimental control, these secondary parameters required consistency throughout the trials.

Unfortunately, if all scenarios appeared to be the same, it would likely support a learning effect. A learning effect can result when participants are asked to perform repetitive tasks or are exposed to consistent environments and as a result they perform better in later trials. For example, if all conflicts involved aircraft at 90 degree angles approaching from the right side, pilots would likely recognize this after only a few runs resulting in hyper-vigilance. Even if each scenario was to be used only twice (for example, once in the high certainty condition and again in the low), there was a chance of participant recognition on the second trial. This was undesirable.

The requirement was to create scenarios which were experimentally consistent while providing a unique scene to the participant with each scenario.

Conflict angle

Two conflict angles were chosen. Approximate conflict geometries were 90 degrees and 60 degrees. The selected angles provided visually discernible stimuli to the pilots. These geometries have shown to be similar for pilot recognition. Mackintosh et al. (2001), only found an effect for obtuse conflict angles. This was the first parameter which was designed to be experimentally consistent, while providing the pilot with a distinct presentation.

Second, with the elimination of right-of-way rules and the requirement that subjects resolve all conflicts, approach geometry from the left and the right were considered experimentally tantamount. Right of way rules are usually in place to support implicit negotiation between aircraft. For example, an aircraft attempting to overtake a second aircraft must yield to that aircraft – thus giving it the right of way. This would be analogous to the downhill skier having the right of way.

These two manipulations provided four (2 x 2) distinct conflict categories (90L, 90R, 60L, 60R). This matrix integrated with the three look-ahead levels (2 min, 4min, 8min) generated twelve distinct trials, a unique traffic scenario for every run.

Traffic density

A previous free flight study (Mackintosh et al., 1998) categorized a high density condition as 15-16 aircraft for the same look-ahead range of 160 nmi. Hoekstra, van Gent and Ruigrok, suggested experiments in free flight utilize double and triple the current European traffic densities (Mackintosh et al., 1998). The traffic for this study was held relatively constant for each scenario. The number of aircraft on screen varied between 14 and 18 at the start of each scenario. That is, 14-18 active aircraft appeared on the CDTI at the start of each run. This number of experimental targets was similar to others investigating free flight, such as Smith Briggs, Knecht, & Hancock (1997). The default setting of the CDTI at start-up was 160 nmi look-ahead. Zooming in or out would decrease or increase the number of aircraft presented to the participant on the monitor. To further prevent recognition, call signs were also changed.

Airspace

Finally, a generic airspace was used to assure that pilots had no prior experience with the experimental airspace. It was believed that if pilots recognized waypoints it might activate a schema associated with that airspace. Since it was unlikely all pilots would exhibit the same behavior, it was decided that generic waypoints (such as WPT-A, WPT-B) would serve as a better control for this study.

With the scenarios constructed to balance control and diversity, a suitable presentation schedule was needed. The one issue scenario fabrication could not address was participant learning during successive trials. A sufficient presentation scheme was needed.

The twelve scenarios were split up into two groups of six (one for each level of certainty). Left and right pairs were separated (group 1: 8min60degLeft group 2: 8min60degRight). In each group a scenario was assigned a letter A – F. The rows from a 6 x 6 Latin square (Box, Hunter, & Hunter1978) were used to create six different presentation orders. High and low certainty blocks were then paired and counterbalanced creating 12 unique trial orders. Six participants received the high certainty condition first, while the other six were first exposed to the low certainty condition.

APPENDIX G

Traffic Display

Cockpit Display of Traffic Information (CDTI)

In order to perform the tasks of separation maintenance and conflict avoidance on the flight deck, the pilots must be able to "see" the surrounding traffic. Pilot awareness of pertinent aircraft including flight parameters (altitude, relative heading, vertical speed), and flight plan is supported on the Cockpit Display of Traffic Information (CDTI)

A prototype CDTI developed at NASA Ames (Johnon, Battiste, Delzell, Holland, Belcher & Jordan, 1997; Johnson, Batiste, & Bochow, 1999) has been in operation since the mid 1990s and used in several studies in addition to their own (DAG-TM Team, 1999). The CDTI (Figure 1) presents the pertinent information about ownship characteristics (altitude, speed, heading, waypoint information, and graphical flight plan similar to a common Visual Navigation Display found on the 747-400), and graphically displays equivalent information regarding the surrounding traffic. The CDTI provided the platform to run and monitor all scenarios for this study.

The display design and symbology maturation were governed by the needs of its three primary functions as stated by the authors. The first two, "basic traffic display" and "conflict alerting logic" (p. 3), were critical for this demonstration.

The display depicts aircraft within a 200 nmi. lateral range and ± 4000 ft vertically. Ownship is portrayed by a filled white chevron with the apex pointing in the direction of travel. Co-altitude aircraft were unfilled chevrons outlined in white. Aircraft below were shown with green outlines and those at higher altitudes than ownship outlined in blue.

When an alert was issued both the intruder and ownship became filled, amber chevrons. An amber line also extended from the points of each aircraft. These lines ended just before the point of intersection. Each line was capped with a small circle representing the minimum separation distance. The touching circles were an indication that a loss of separation was predicted. Also appearing in the lower left-hand corner of the screen was the word "ALERT" and beneath that "LOS 0:00" with the appropriate time indicated. Both were in amber (Figure 1). Finally, a single chime was issued making the alert multi-modal.

APPENDIX H

Free Flight Brief

Thank you for supporting this experiment with you participation.

Today you will be a participant in a free flight study. The study will be performed in a motionless, part task environment utilizing desktop PCs. You will receive a briefing and training which should take 1-2 hours. The experimental runs will be administered in two sessions with a break between the sessions. A short debrief will follow the completion of the trials.

About the experiment

Free flight is a proposed air traffic management paradigm designed to increase the flexibility of the National Airspace System (NAS). The distribution of separation responsibility between ATC and the flight deck is thought to be a pillar of this design. With the responsibility of separation, benefits emerge such as user preferred routing. This would allow pilots to select shorter, more direct routes and/or take advantage of wind conditions. The inefficiency of strict adherence to prescribed airways would be replaced with flight specific augmentations.

In order to perform tasks such as self-separation and conflict resolution, additional tools will need to be provided. One such tool is a cockpit display of traffic information (CDTI). This will provide the flight deck with the pertinent information of the surrounding traffic such as altitude, call sign and speed. Another function critical to the safe operation of this new air traffic system is a conflict probe that will alert the flight deck if a loss of separation with another aircraft (intruder) is predicted.

Conflict prediction is an inexact science. Winds, aircraft performance and navigation, and a host of other elements contribute to discrepancies in any prediction. Uncertainty in alerting can lead to false alarms or missed alerts. Unfortunately, in the design of detection systems these error types cannot be reduced simultaneously.

Therefore, this experiment was designed to examine the impact of two different levels of certainty. A high alerting level condition and low level condition will be administered. A system threshold set such that it only activates when it is very sure of a conflict (> 99%) is susceptible to missing a true conflict. Conversely, while a lower alerting threshold (75%) is less likely to miss an intruder, the tendency for false alerts emerges. More about this will be covered in the training.

Please feel free to ask any questions along the way. A formal break will be given between sessions, but any time you wish to stop, please inform me.

Thanks again for your participation.

Paul Picciano

APPENDIX I

Participant Profile

DATE__________________ Subject #__________________

Circle: CAPTAIN - FIRST OFFICER

Age:_________

Please Estimate your Flight Hours:

Total Flight Hours _________________

Flight Hours on Glass Cockpit_________________

Current/Most Recent Aircraft__________________

If current aircraft not glass, when did you last fly glass?____________________

Total Flight Hours on Current/Most Recent Aircraft___________________

Typical Trip Destinations/ Number per month_____________________________

____

____

Previous knowledge of Free Flight: (circle one)

What?

I've heard of it.

I know a little about it.

Actively involved in the FF community.

APPENDIX J

Consent Form

Paul Picciano, Kevin Corker, Ph. D., Kevin Jordan, Ph. D. , Sandy Lozito

AGREEMENT TO PARTICIPATE IN RESEARCH

Time vs. Certainty: Pilot preference and behavior under varied conflict alerting schemes in a free flight envrinoment.

I have been asked to participate in a research study investigating the impact of varying conflict-alerting schemes on pilot preference and behavior. I will participate as a pilot at a desktop computer simulator configuration at NASA Ames. The simulation will be composed of two desktop computers, communication equipment and video recording apparatus. Since this is a stationary simulator environment, no significant risks are anticipated in performing this study. The purpose of this research is to explore issues of free flight. The standard pay for participation as well as travel reimbursements will be allotted. The results of this study may be published but individual participants will remain unidentified. General questions about this experiment can be answered by the principal investigator Paul Picciano, Email: ppicciano@mail.arc.nasa.gov, Phone: (650) 604-0050. Questions and complaints about the research may also be addressed to the Human Factors/Ergonomics Department Chair Kevin M. Corker, Ph. D., Email:kcorker@email.sjsu.edu, Phone: (408) 924-3988. Questions or complaints about research, subjects' rights, or research related injury may be presented to Nabil Ibrahim, Ph. D., Associate Academic Vice President for Graduate Studies and Research, at (408) 924-2480. No service of any kind will be lost or jeopardized if at anytime you choose to terminate your participation in the study. Your signature implies your voluntary consent to take part in the study. However, you are free to withdraw at anytime without prejudice to your relations with San Jose State University or NASA. I will receive a signed and dated copy of the consent form.

Initial _______

_____________________________ ______________

Participant's Signature Date

_____________________________ ______________

Investigator's Signature Date

Addendum

Recruitment for this study will be handled by Ratheon, a contractor at the NASA Ames Research Center. This is a standard protocol where confidentiality and the property of the data are ensured.

San José State
UNIVERSITY

Office of the Academic Vice President

Associate Vice President Graduate Studies and Research

One Washington Square
San Jose, CA 95192-0025
Voice: 408-283-7500
Fax: 408-924-2477
E-mail: gstudies@wahoo.sjsu.edu
http://www.sjsu.edu

To: Paul Picciano
NASA Ames Research Center
M/S 262-4
Moffett Field, CA 94035

From: Nabil Ibrahim,
AVP, Graduate ~~Studies &~~ Research

Date: March 22, 2002

The Human Subjects-Institutional Review Board has approved your request to use human subjects in the study entitled:

"Time vs. Certainty: Pilot preference and behavior under varied conflict alerting schemes in a free flight environment."

This approval is contingent upon the subjects participating in your research project being appropriately protected from risk. This includes the protection of the anonymity of the subjects' identity when they participate in your research project, and with regard to any and all data that may be collected from the subjects. The approval includes continued monitoring of your research by the Board to assure that the subjects are being adequately and properly protected from such risks. If at any time a subject becomes injured or complains of injury, you must notify Nabil Ibrahim, Ph.D. immediately. Injury includes but is not limited to bodily harm, psychological trauma, and release of potentially damaging personal information. This approval for the human subjects portion of your project is in effect for one year, and data collection beyond March 21, 2003 requires an extension request.

Please also be advised that all subjects need to be fully informed and aware that their participation in your research project is voluntary, and that he or she may withdraw from the project at any time. Further, a subject's participation, refusal to participate, or withdrawal will not affect any services that the subject is receiving or will receive at the institution in which the research is being conducted.

If you have any questions, please contact me at (408) 924-2480.