



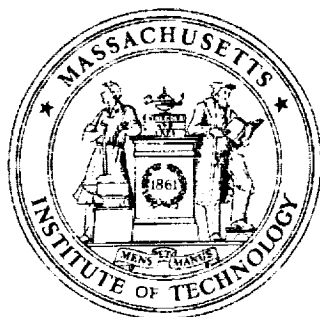
MIT AERONAUTICAL SYSTEMS LABORATORY

**EXPERIMENTAL STUDY OF COLLISION DETECTION
SCHEMA USED BY PILOTS DURING CLOSELY SPACED
PARALLEL APPROACHES**

Amy R. Pritchett and R. John Hansman

Aeronautical Systems Laboratory
Department of Aeronautics & Astronautics
Massachusetts Institute of Technology
Cambridge, Massachusetts USA

January 1996
ASL-96-1



**DEPARTMENT OF
AERONAUTICS AND ASTRONAUTICS**
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS 02139-4307



MIT AERONAUTICAL SYSTEMS LABORATORY

**EXPERIMENTAL STUDY OF COLLISION DETECTION
SCHEMA USED BY PILOTS DURING CLOSELY SPACED
PARALLEL APPROACHES**

**Amy R. Pritchett
R. John Hansman**

**Aeronautical Systems Laboratory
Department of Aeronautics & Astronautics
Massachusetts Institute of Technology
Cambridge, Massachusetts USA**

January 1996

ASL-96-1

EXPERIMENTAL STUDY OF COLLISION DETECTION SCHEMA USED BY PILOTS DURING CLOSELY SPACED PARALLEL APPROACHES

**Amy R. Pritchett
R. John Hansman**

**Aeronautical Systems Laboratory
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, Massachusetts USA**

5 January 1996

ASL-96-1

Abstract

An experimental flight simulator study was conducted to examine the mental alerting logic and thresholds used by subjects to issue an alert and execute an avoidance maneuver. Subjects flew a series of autopilot landing approaches with traffic on a closely-spaced parallel approach; during some runs, the traffic would deviate towards the subject and the subject was to indicate the point when they recognized the potential traffic conflict, and then indicate a direction of flight for an avoidance maneuver. A variety of subjects, including graduate students, general aviation pilots and airline pilots, were tested. Five traffic displays were evaluated, with a moving map TCAS-type traffic display as a baseline. A side-task created both high and low workload situations.

Subjects appeared to use the lateral deviation of the intruder aircraft from its approach path as the criteria for an alert regardless of the display available. However, with displays showing heading and/or trend information, their alerting thresholds were significantly lowered. This type of range-only schema still resulted in many near misses, as a high convergence rate was often established by the time of the subject's alert. Therefore, the properties of the intruder's trajectory had the greatest effect on the resultant near miss rate; no display system reliably caused alerts timely enough for certain collision avoidance. Subjects' performance dropped significantly on a side-task while they analyzed the need for an alert, showing alert generation can be a high workload situation at critical times. No variation was found between subjects with and without piloting experience.

These results suggest the design of automatic alerting systems should take into account the range-type alerting schema used by the human, such that the rationale for the automatic alert should be obvious to, and trusted by, the operator. Although careful display design may help generate pilot/automation trust, issues such as user non-conformance to automatically generated commands can remain a possibility.

Acknowledgments

This work was supported by the National Aeronautics and Space Administration/ Ames Research Center under grant NAG 2-716.

The authors would like to acknowledge the following persons for their contributions to this study:

Kevin Corker, Sandy Lozito, and Trent Thrush of NASA Ames Research Centre for their review of the experiment design.

The students, flight instructors and airline pilots who donated time to fly the simulator as subjects.

Table of Contents

ACKNOWLEDGMENTS.....	1
TABLE OF CONTENTS.....	2
FIGURES.....	3
TABLES.....	4
1. INTRODUCTION AND MOTIVATION.....	5
1.1 CLOSELY SPACED PARALLEL APPROACHES.....	5
1.2 PREVIOUS EXPERIMENTAL RESULTS.....	5
1.3 EXPERIMENT OBJECTIVES.....	13
2. EXPERIMENT DESIGN.....	14
2.1 TEST MATRIX.....	14
2.1.1 Displays.....	14
2.1.2 Workload Levels.....	14
2.1.3 Scenarios.....	16
2.1.4 Test Matrix.....	17
2.2 EXPERIMENTAL PROCEDURE.....	17
2.3 SIMULATOR SETUP.....	18
2.4 SUBJECTS.....	19
2.5 MEASUREMENTS.....	19
3. EXPERIMENT RESULTS.....	20
3.1 OVERALL PERFORMANCE.....	20
3.2 DISPLAY EFFECTS.....	22
3.3 WORKLOAD EFFECTS.....	25
3.4 COLLISION GEOMETRY EFFECTS.....	27
3.4.1 Differences in Performance Between the Different Scenarios.....	27
3.4.2 Differences in Performance Based on Aircraft Convergence Rate.....	29
3.4.3 Differences in Performance Based on Turn Rate.....	32
3.4.4 Differences in Performance Based on Intruder Vertical Speed.....	35
3.5 VARIANCE WITH SUBJECT CHARACTERISTICS.....	37
4. CONCLUSIONS.....	39
APPENDIX A. DEBRIEFING FORM.....	40
REFERENCES.....	41

Figures

FIGURE 1. CARTOON OF CLOSELY SPACED PARALLEL APPROACHES.....	5
FIGURE 2. BASELINE TRAFFIC DISPLAY, SHOWING TRAFFIC ON PARALLEL APPROACH.....	6
FIGURE 3. ENHANCED EHSI TRAFFIC DISPLAY.....	7
FIGURE 4. TRAFFIC DISPLAY ON PFD	8
FIGURE 5. SUBJECTIVE AHP RATINGS OF THE DISPLAYS IN THE BASELINE EXPERIMENT.....	9
FIGURE 6. PERCENTAGE OF APPROACHES RESULTING IN A SPURIOUS AVOIDANCE MANEUVER.....	9
FIGURE 7. HISTOGRAM OF PILOT REACTION TIMES TO ALERTS WITH EACH DISPLAY.....	10
FIGURE 8. PERCENTAGE OF APPROACHES RESULTING IN LOW AIRCRAFT SEPARATION WITH EACH DISPLAY.....	10
FIGURE 9. EVALUATION OF PILOT CONFORMANCE TO COMMANDED TCAS MANEUVER	11
FIGURE 10. CONFORMANCE RATE WITH EACH DISPLAY.....	11
FIGURE 11. RANGE VERSUS RANGE-RATE PLOT.....	12
FIGURE 12. BASELINE DISPLAY SHOWING OTHER AIRCRAFT POSITION.....	15
FIGURE 13. FEDUCIAL MARK DISPLAY WITH PARALLEL APPROACH COURSE SHOWN.....	15
FIGURE 14. HEADING DISPLAY WITH HEADING INFORMATION ADDED	15
FIGURE 15. DISPLAYS WITH PROJECTED POSITION (15 SECONDS)	15
FIGURE 16. EXPERIMENT TEST MATRIX.....	17
FIGURE 17. MANEUVER SELECTION SCREEN	18
FIGURE 18. RANGE VERSUS RANGE RATE PLOT WHEN SUBJECTS GENERATED AN ALERT	21
FIGURE 19. HISTOGRAM OF AIRCRAFT LATERAL SEPARATION WHEN SUBJECTS GENERATED ALERT	21
FIGURE 20. AVERAGE REACTION TIME WITH EACH DISPLAY.....	23
FIGURE 21. AVERAGE INTRUDER DEVIATION FROM FEDUCIAL MARK AT SUBJECT ALERT TIMES WITH EACH DISPLAY	24
FIGURE 22. FREQUENCY OF NEAR-MISSES / COLLISIONS WITH EACH DISPLAY	25
FIGURE 23. FREQUENCY OF NEAR-MISSES / COLLISIONS WITH DIFFERENT SUBJECT WORKLOADS.....	26
FIGURE 24. MSE BANK FOR DIFFERENT SUBJECT WORKLOADS.....	26
FIGURE 25. AVERAGE REACTION TIME WITH THE DIFFERENT SCENARIO TYPES.....	27
FIGURE 26. AVERAGE INTRUDER DEVIATION FROM APPROACH PATH WITH THE DIFFERENT SCENARIO TYPES.....	28
FIGURE 27. FREQUENCY OF NEAR-MISSES / COLLISIONS WITH THE DIFFERENT SCENARIO TYPES.....	29
FIGURE 28. HISTOGRAMS OF REACTION TIME TO HIGH AND LOW CONVERGENCE INTRUSIONS	30
FIGURE 29. HISTOGRAMS OF AIRCRAFT LATERAL SEPARATION AT SUBJECTS' REACTIONS	30
FIGURE 30. FREQUENCY OF AVOIDANCE MANEUVER SELECTION WITH INTER-AIRCRAFT CONVERGENCE RATE	31
FIGURE 31. FREQUENCY OF COLLISIONS WITH HIGH AND LOW CONVERGENCE INTRUSIONS.....	32
FIGURE 32. AVERAGE REACTION TIMES WITH DIFFERENT CONVERGENCE AND TURN RATES.....	33
FIGURE 33. HISTOGRAM OF INTRUDER LATERAL DEVIATION FROM NOMINAL APPROACH PATH AT SUBJECT REACTIONS WITH DIFFERENT CONVERGENCE AND TURN RATES.....	33
FIGURE 34. SCATTER PLOT COMPARING REACTION TIME AND TIME REMAINING WITH DIFFERENT CONVERGENCE AND TURN RATES.....	34
FIGURE 35. FREQUENCY OF NEAR-MISSES / COLLISIONS WITH DIFFERENT CONVERGENCE AND TURN RATES.....	34
FIGURE 36. INTRUDER RELATIVE ALTITUDE AT SUBJECT REACTION TIME, BY DIFFERENT COMMANDED VERTICAL SPEEDS.....	35
FIGURE 37. INTRUDER VERTICAL SPEED AT SUBJECT REACTION TIME, BY DIFFERENT COMMANDED VERTICAL SPEEDS.....	36
FIGURE 38. FREQUENCY OF AVOIDANCE MANEUVER SELECTION, BY DIFFERENT COMMANDED VERTICAL SPEEDS.....	37
FIGURE 39. AVERAGE REACTION TIMES OF DIFFERENT AGE GROUPS.....	38
FIGURE 40. MSE BANK DURING WINGS-LEVELING TASK OF DIFFERENT AGE GROUPS.....	38

Tables

TABLE 1. RELATIVE PERFORMANCE OF ENHANCED TRAFFIC DISPLAYS IN BASELINE EXPERIMENT	8
TABLE 2. TIME FROM INTRUDER LEAVING OWN APPROACH PATH TO THE CROSSING OF THE SUBJECT'S APPROACH PATH (INTRUDER SPEED 140-180 KNOTS).....	16
TABLE 3. FREQUENCY AND TYPES OF ALERTING DECISIONS.....	20
TABLE 4. FREQUENCY OF AVOIDANCE MANEUVERS SELECTION.....	22
TABLE 5. DISPLAYS TESTED IN THE EXPERIMENT.....	22
TABLE 6. CORRECTNESS OF SUBJECT ALERT GENERATION	23
TABLE 7. CORRECTNESS OF THE SUBJECTS' ALERTS TO THE HIGH AND LOW CONVERGENCE INTRUSIONS	29
TABLE 8. TOTAL BLUNDER DURATION FOR DIFFERENT CONVERGENCE & TURN RATES	32

1. Introduction and Motivation

1.1 Closely Spaced Parallel Approaches

Several major airports around the United States have, or plan to have, closely-spaced parallel runways, allowing aircraft to land on both runways simultaneously. This requires the aircraft to fly close to each other during their parallel landing approaches, as shown in Figure 1. During Visual Meteorological Conditions (VMC), the responsibility for collision avoidance is given to the pilots, who are to maintain visual contact with each other. However, during Instrument Meteorological Conditions (IMC), the responsibility is currently given to air traffic controllers. Current technology limits independent parallel approaches to runways spaced 4300 feet or more apart (3000' feet with the Precision Runway Monitor, which uses specialized ground based radar and a dedicated controller, in place at Memphis and Raleigh-Durham airports); the use of new technologies to reduce this minimum separation would allow airports to effectively maintain their high VMC capacity in IMC.

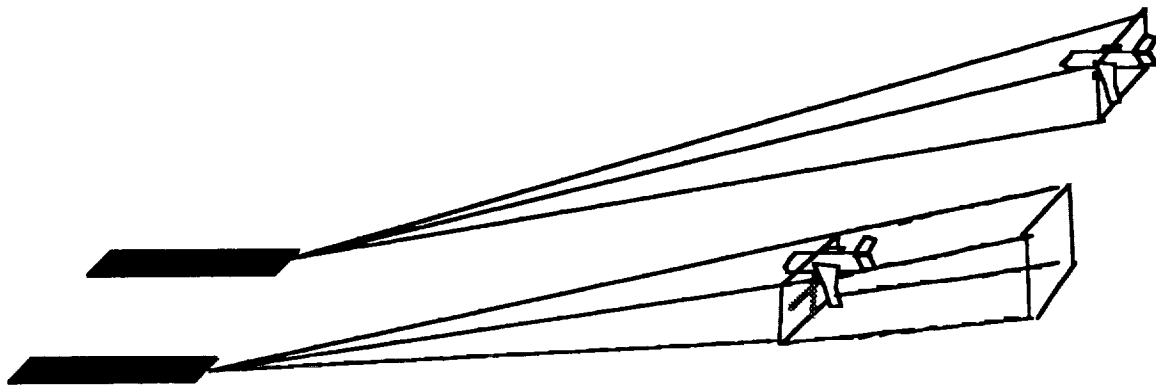


Figure 1. Cartoon of Closely Spaced Parallel Approaches

Several studies are in process studying the various elements of an airborne collision avoidance required for closely spaced parallel approaches. A great deal of attention is especially being paid to the cockpit traffic displays required for pilot situation awareness. It is hoped that features of traffic displays can be identified that make the pilot more comfortable with parallel approach operations, help pilots react to potential conflicts sooner and allow pilots to better execute avoidance maneuvers which maintain adequate separation from other aircraft.

1.2 Previous Experimental Results

A baseline experiment has already been completed which examined both procedural and display issues [Pritchett & Hansman, In Progress]. This experiment provided active airline pilots with four different displays: as a baseline, a representation of a traffic display integrated onto the Electronic Horizontal Situation Indicator (EHSI), such as is currently provided as part of TCAS II, shown in Figure 2; enhancements to this EHSI display giving expanded scales and reference indications of the parallel approach position, shown in Figure 3; a back-to-front view of relative lateral and vertical position of the traffic on the Primary Flight Display (PFD) shown with the baseline EHSI traffic display, shown in Figure 4; and finally, the combination of the Enhanced EHSI and PFD displays.

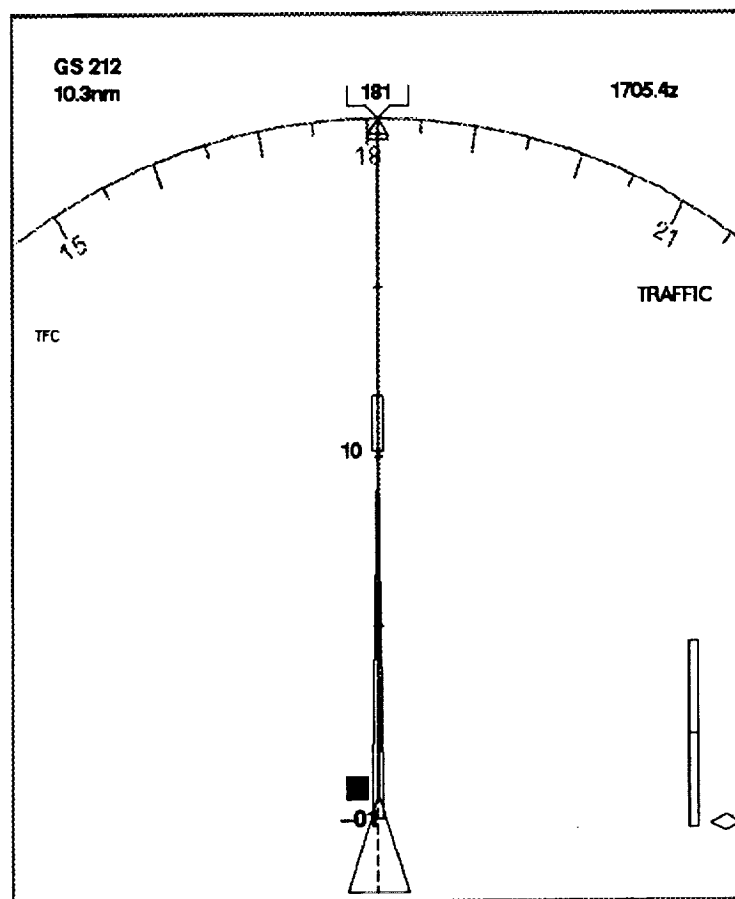


Figure 2. Baseline Traffic Display, Showing Traffic on Parallel Approach

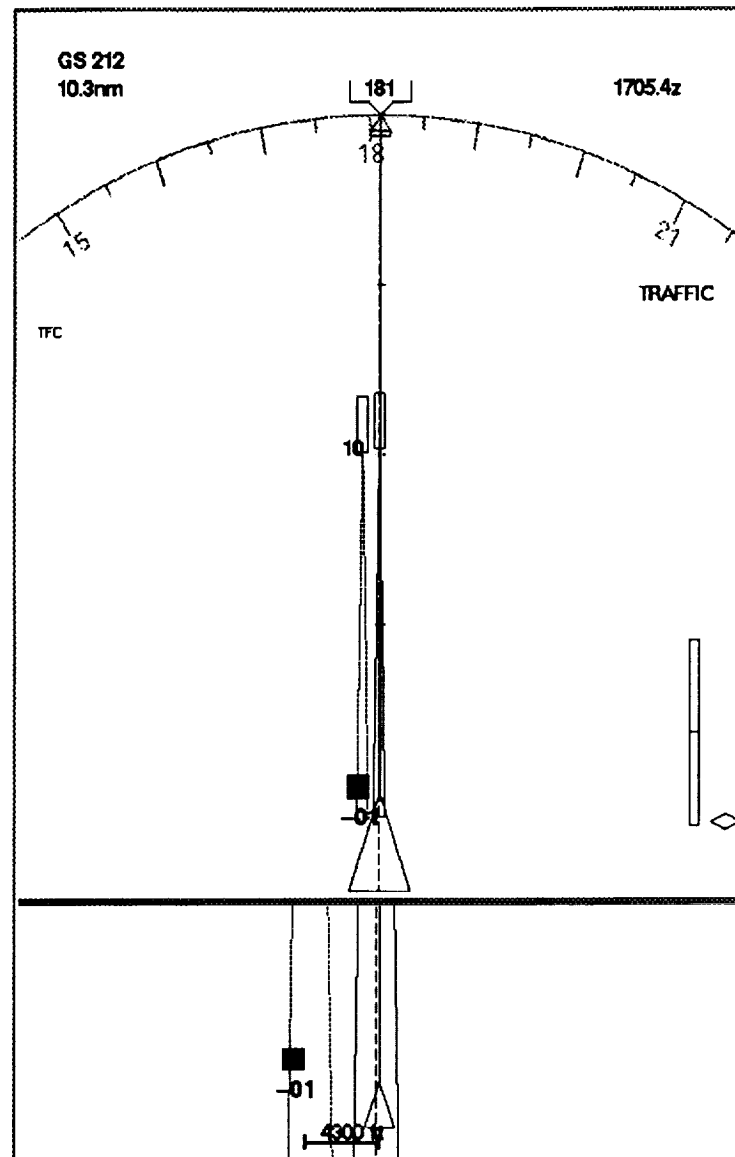


Figure 3. Enhanced EHSI Traffic Display

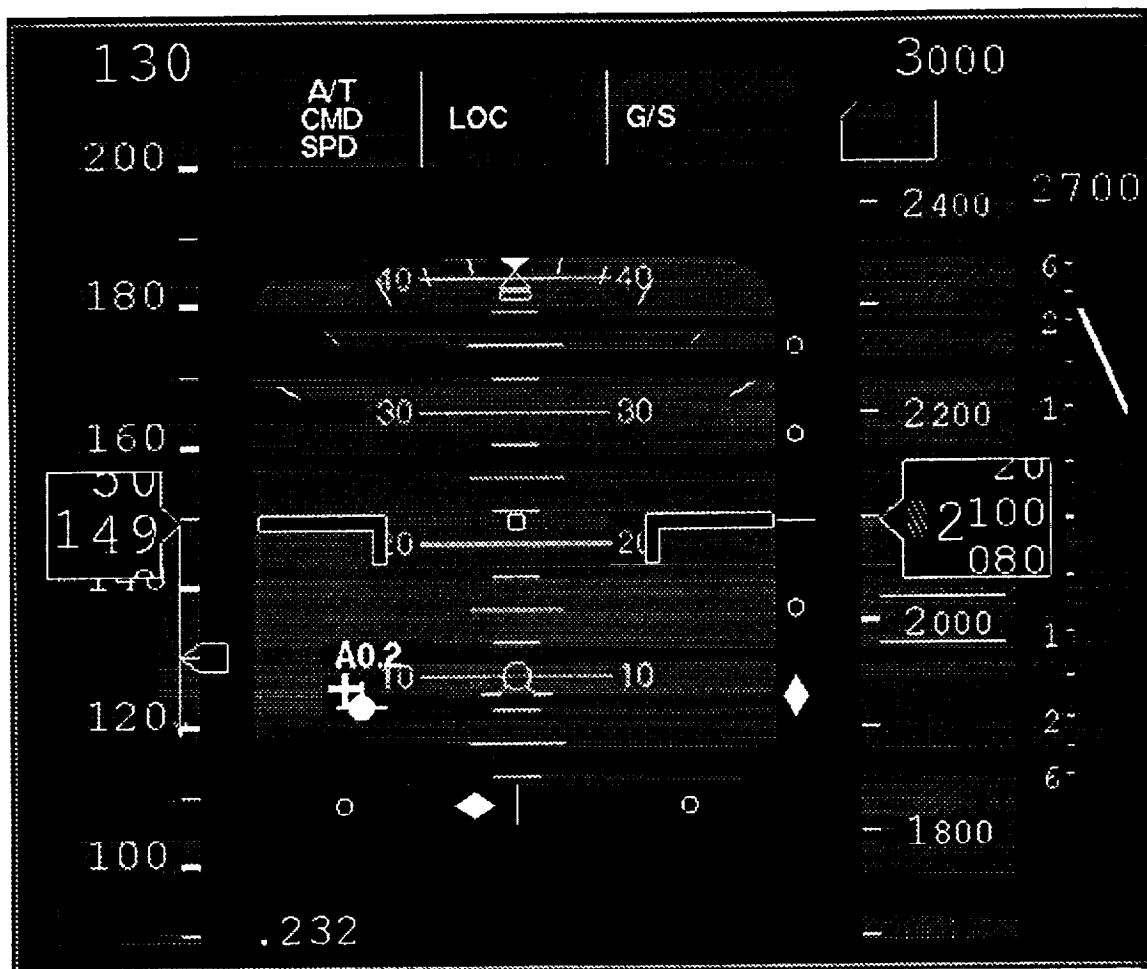


Figure 4. Traffic Display on PFD

The results of this baseline experiment provide conflicting evidence about the benefits of these enhanced cockpit displays, as summarized in Table 1 and detailed below.

Benefits	No Improvement	Drawbacks
Higher Pilot Opinions	Reaction Time	More Near Misses / Collisions
Less Spurious Go-Arounds		Less Conformance to TCAS Commanded Maneuvers

Table 1. Relative Performance of Enhanced Traffic Displays in Baseline Experiment

The enhanced EHSI and PFD displays were greatly favoured by the pilots over the current TCAS display. Paired-comparison ratings were made by pilots after the experiment, and the results were combined using the Analytic Hierarchy Process (AHP) [Method described in Yang & Hansman, 1995]. These AHP ratings were normalized to sum to one, and their relative percentages are shown in Figure 5. The pilots' comparative approval of the ratings can be found by finding the multiplicative difference between them. Having the combination of the Enhanced EHSI and PFD displays was considered 11 times better than the TCAS II type baseline display (55% / 5%). The enhanced EHSI alone was considered about five times better than the baseline TCAS II display (27% / 5%), and about twice as

good as the PFD and baseline display combination (27% / 13%). Finally, the combination of the PFD display and the baseline was considered almost three times better than the baseline display alone (13% / 5%).

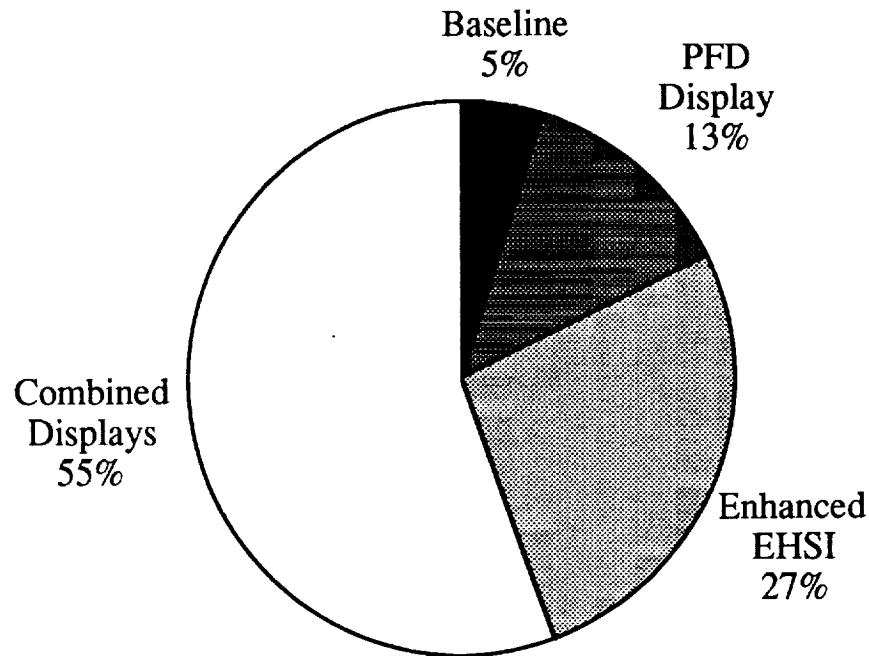


Figure 5. Subjective AHP Ratings of the Displays in the Baseline Experiment

In addition, the enhanced displays made pilots more comfortable with close parallel approaches, as shown by the drop in 'spurious' avoidance maneuvers flown before the parallel traffic deviated from its acceptable approach course and before the TCAS alerts (when available) were issued. As shown in Figure 6, the drop is very significant ($p < 0.01$) from the baseline display to the enhanced EHSI display.

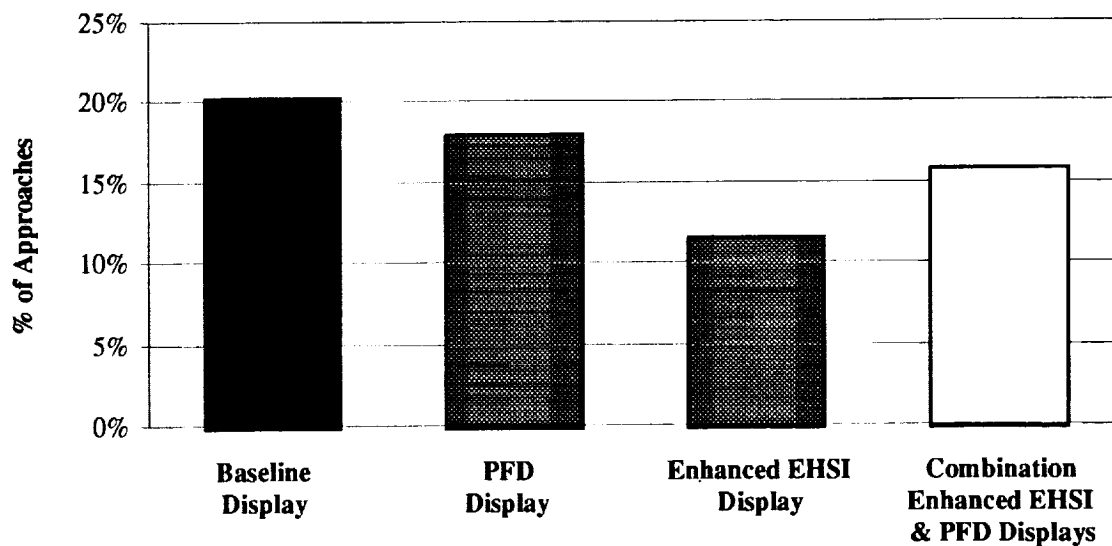


Figure 6. Percentage of Approaches Resulting in a Spurious Avoidance Maneuver

Although a small trend of reduced reaction time to automatically generated alerts was found when the pilots were using the enhanced displays, these differences can not be judged statistically significant. The distributions of reaction time for each display are essentially identical, as can be seen from Figure 7. It should be noted that these reaction times, given the one-near-miss-per-approach nature of this experiment, probably do not represent those found from pilots in real operations, where potential collisions are rare.

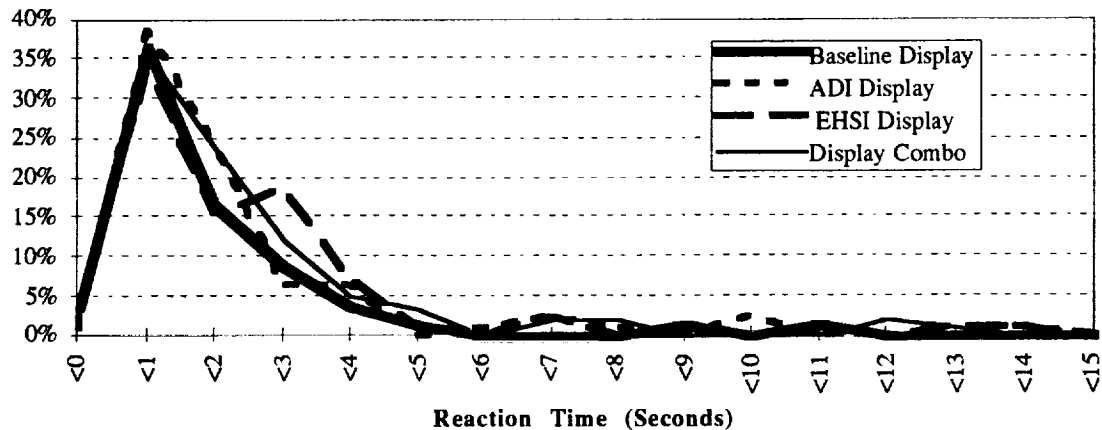


Figure 7. Histogram of Pilot Reaction Times to Alerts with Each Display

For collision avoidance systems, the dominant performance metric is the frequency of low aircraft separation incidents, defined in this baseline experiment as collisions (miss distances less than 500 feet) and near misses (miss distances less than 1000 feet). Performance with the enhanced displays was degraded, as shown by the significantly higher incident rates in Figure 8. The difference between these incident rates for the baseline display and the enhanced EHSI display is significant ($p < 0.05$).

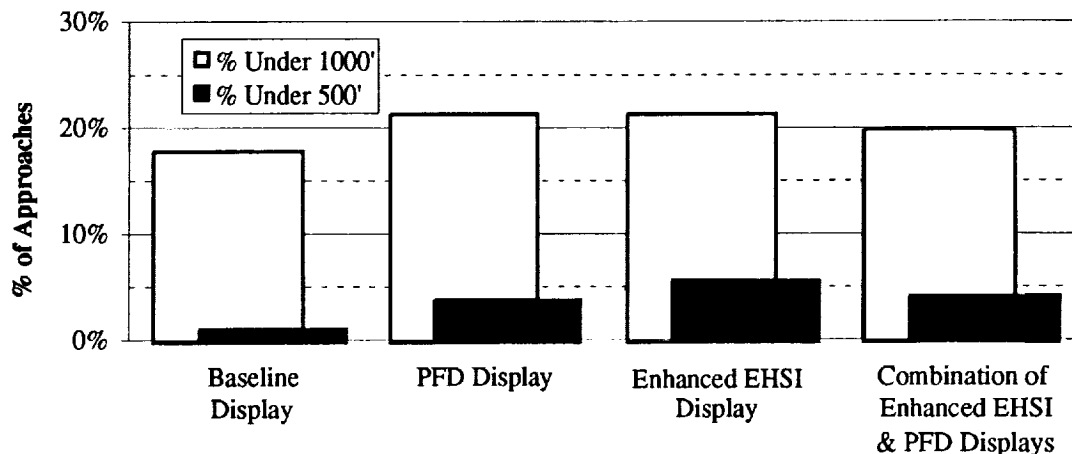


Figure 8. Percentage of Approaches Resulting in Low Aircraft Separation with Each Display

Finally, pilots in the baseline experiment had, in most of the runs, a TCAS II alerting system available which alerted the pilots to a possible traffic hazard and displayed a vertical (climb or descend) avoidance maneuver on the PFD. This system was designed assuming a five second pilot reaction time, followed by a 0.25g pull-up or push-over maneuver to the commanded pitch attitude. In this analysis, any maneuver that met or exceeded this criteria was considered to be in conformance with the TCAS command, as illustrated in Figure 9 for a commanded climb.

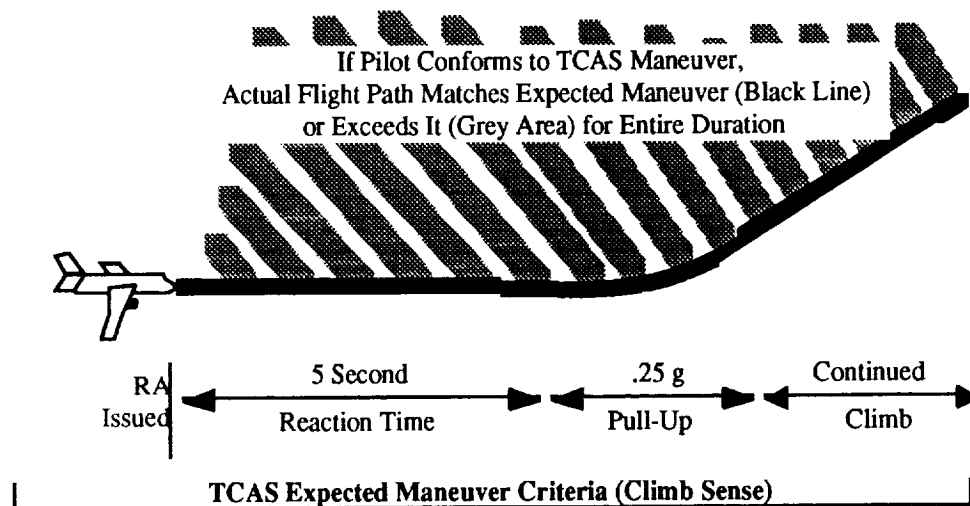


Figure 9. Evaluation of Pilot Conformance to Commanded TCAS Maneuver

Pilot conformance to these avoidance maneuvers was found to decrease significantly with the enhanced displays, as shown in Figure 10. The difference between the conformance with the baseline display and with the enhanced EHSI display is very significant ($p < 0.01$); the difference between the baseline and PFD display is significant ($p < 0.05$).

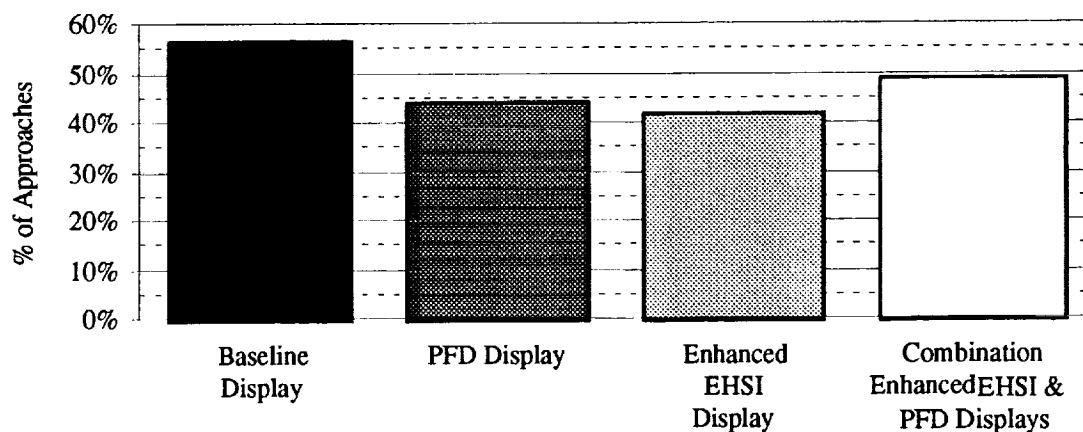


Figure 10. Conformance Rate with Each Display

In summary, these enhanced traffic displays were designed with three goals in mind: that they increase pilot confidence in flying close parallel approaches; that they reduce reaction time; and that they allow the pilot to maximize the separation between their aircraft and the intruding aircraft. The displays tested in the baseline experiment helped the first goal, caused no noticeable difference in the second, and were detrimental towards the third.

Further analysis suggested that these problems may stem from the pilots' use of a different and less effective alerting algorithm for deciding when to generate alerts. For example, pilots, when not presented with TCAS - type alerts, flew a maneuver matching what TCAS would have commanded only 25% of the time, suggesting their instinctive reaction is not always what the TCAS system commands.

Plotting the position of the intruder relative to the subject at the point where the subject reacted shows their reactions were very scattered. However, it also appears that the across-track deviation of the intruding aircraft appears to be a major determinant in the decision to react, a conclusion also supported by pilot comments about their decisions to alert.

An alert generation logic based on intruder lateral deviation is comparable to the Non-Transgression Zone (NTZ) type alerting algorithm used by the Parallel Runway Monitoring (PRM) system [Shank & Hollister, 1994]. This type of alerting algorithm has been shown to be ineffective for two reasons: it can generate a false alarm when the parallel traffic oscillate around their localizer course during a normal approach [Owen, 1993], and it may not generate an alert until the intruding aircraft has already established a high rate of convergence.

The TCAS alerts shown to the pilots during the baseline experiment were generated by a different algorithm, which uses inter-aircraft range and convergence rate (range rate). Plotting the range versus range-rate at both the times when the TCAS alerts were generated and when the pilots generated an alert (when they were not shown TCAS alerts), we see that the pilots' reactions do not appear to take into account range-rate, as shown in Figure 11. Instead, as mentioned before, the pilots seemed to predominantly consider range to the other aircraft.

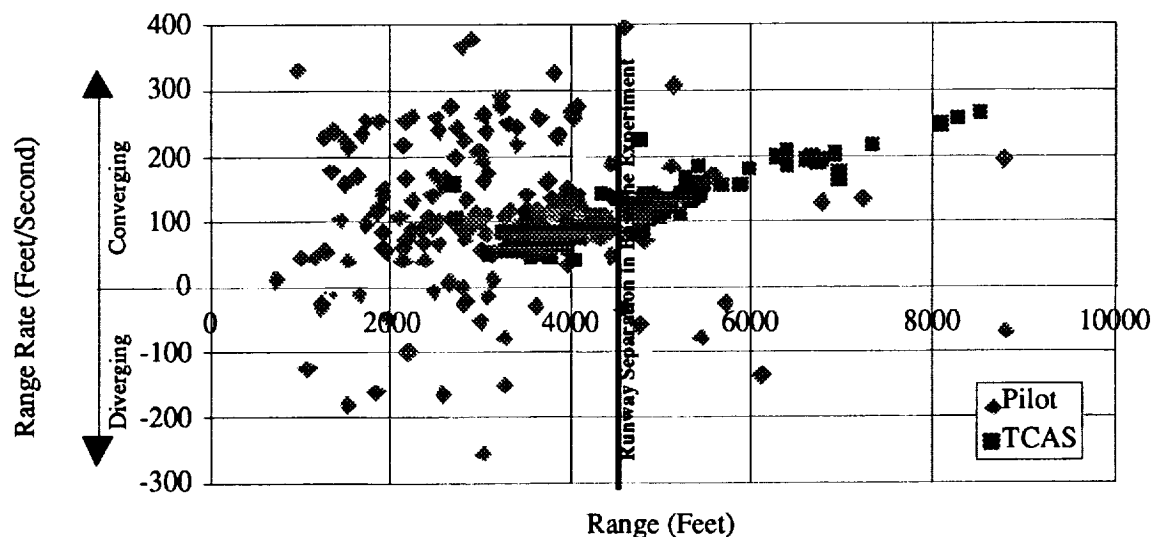


Figure 11. Range versus Range-Rate Plot
Conditions at Time of TCAS Alerts and Pilot Generated Alerts

The enhanced displays tested in the baseline experiment provided pilots with a fiducial marker indicating the cross-track position of a normal approach. All pilots indicated they liked this feature; some commented that it freed them from monitoring the convergence rate

of the other aircraft. Therefore, this feature may have unintentionally encouraged a range-only alerting logic.

It was hypothesized that the traffic display features can, and should, support a more sophisticated mental model for pilots to use in generating alerts and commencing avoidance maneuvers. This should provide for better pilot confidence in, and following of, automatically displayed avoidance maneuvers (when available), and reduce erroneous pilot reactions.

1.3 Experiment Objectives

This experiment has the following three objectives:

- 1) Provide a preliminary study of how the display features of a cockpit traffic display affect a person's mental 'alert generation logic', used to assess when an avoidance maneuver is necessary and what the avoidance maneuver should be,
- 2) Ascertain how display features affect a user's ability to detect a conflict, and
- 3) Test the effect of subject workload on this ability.

2. Experiment Design

2.1 Test Matrix

The test matrix for this experiment was three dimensional, varying displays, workload levels and traffic conflict scenarios as detailed below.

2.1.1 Displays

Five displays were tested. All were based on a moving map/ EHSI type display, with a top-down view, heading-up orientation, iconic presentation of the other aircrafts' positions and a text presentation of the other aircrafts' altitudes. All features of the traffic display were updated once per second, an update rate feasible with current technology.

- *Baseline Display*: emulated the current TCAS display, as shown in Figure 2a.
- *Fecudcial Mark Display*: added the reference indication of the parallel approach path, emulating the enhanced EHSI display tested in the baseline experiment as shown in Figure 2b.
- *Heading Display*: added a graphic indication of the other aircrafts' heading as shown in Figure 2c.
- *Noisy Projection Display*: added a graphic indication of heading rate and projected position within the next 15 seconds, as shown in Figure 2d; the position projection was based on the noisy measurement of the other aircrafts' bank, using bank variations typical of those found during localizer tracking,
- *Smooth Projection Display*: added a graphic indication of heading rate and projected position within the next 15 seconds, as shown in Figure 2d; the position projection used exact knowledge of heading rate to give a more smooth projection.

2.1.2 Workload Levels

The subjects were told, during briefing, their primary task was to keep their wings level despite turbulence, using a side-stick. To do this, an artificial horizon was available to them, drawn approximately three inches away from the edge of the traffic display. For an additional incentive, a prize was offered to the subject whose bank deviation from level was the smallest, averaged over all data runs.

The turbulence was set to two different levels: in one case, it required almost all of the subjects' attention, while in the other case it required checking, on average, only once every two seconds. The subjects were not briefed on these qualities. The turbulence in bank was provided by a Markov model, with the frequency of state changes, and the probability of state changes, set for both the High and Low workload cases through preliminary subject runs.

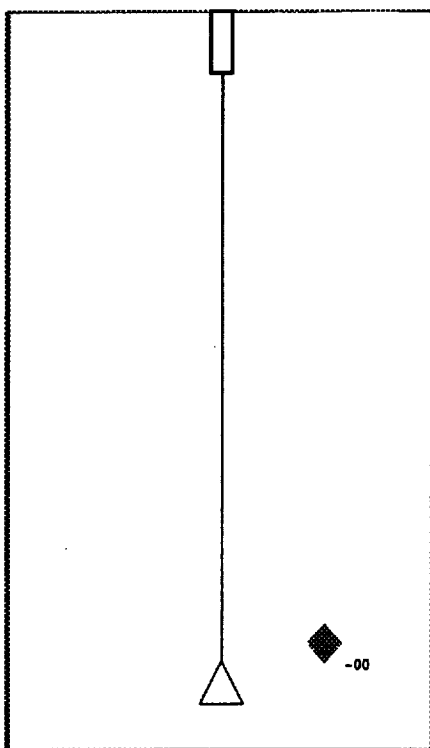


Figure 12. Baseline Display Showing Other Aircraft Position

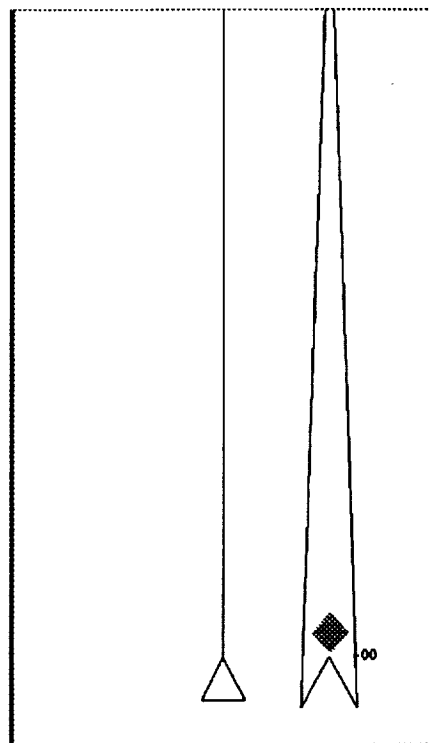


Figure 13. Feducial Mark Display with Parallel Approach Course Shown

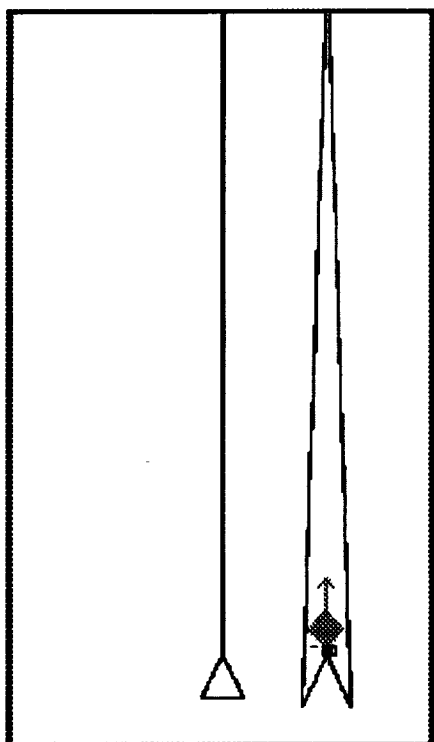


Figure 14. Heading Display with Heading Information Added

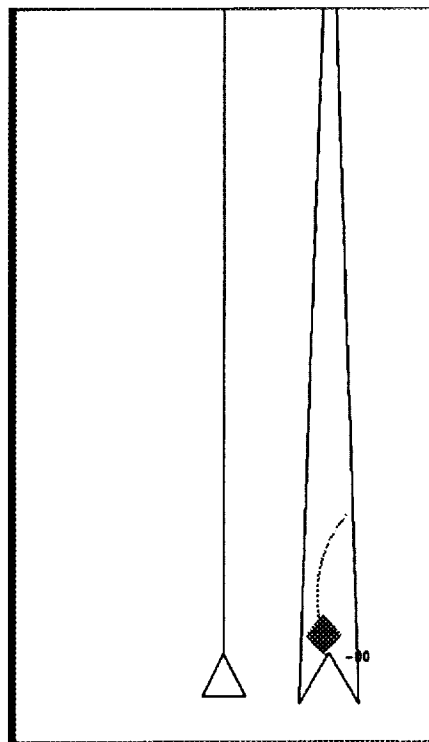


Figure 15. Displays with Projected Position (15 Seconds)

2.1.3 Scenarios

Four scenarios were flown, in random order, within each test block. They were:

Missed Intercept: The parallel traffic does not capture its parallel approach course but continues through its localizer intercept on a straight line collision course.

Hazardous Blunder: The parallel traffic joins its own approach course, but at a random time during the approach turns towards the subject, establishing a collision trajectory.

Less Hazardous Blunder: The parallel traffic joins its approach course, but at a random time during the approach blunders toward the subject, establishing a trajectory that will have it pass at least 1000 feet away from the subject.

Safe Approach: The parallel traffic flies its own approach normally, without deviating.

The intercept angle of the intruding aircraft was picked to be high in one half of the runs with each display (45°) and to be low in the remaining one half of the runs (15°). Likewise, the turn rate of the intruder was set to a high value in one half of the runs ($4.5^\circ/\text{second}$) and low in the remaining half ($1.5^\circ/\text{second}$).

The intruder's speed was randomly selected at the beginning of each experiment run using a uniform distribution between 140 knots (the subject's speed) and 180 knots.

All cases represented parallel approaches 2000 feet apart. Table 2 shows the length of time from the moment when the intruder leaves its own approach path to the time when the intruder crosses the subject's approach path. In the case where the intruder never intercepts its own approach path, this measurement starts when the intruder crosses the centerline of its approach path; in the cases where the intruder does intercept its approach, this measurement starts when the intruder first starts its turn towards the subject. Although this measurement is not necessarily the time to point of closest approach, it can be considered a representation of the overall time-scale of the subjects' task.

		Intercept Angle = 15°	Intercept Angle = 45°
Intruder Never Intercepts Own Approach		32.7 - 25.4 Seconds	12.0 - 9.3 Seconds
Intruder Intercepts Own Approach, Then Deviates	Turn Rate = $1.5^\circ/\text{Sec}$	37.7 - 30.4 Seconds	26.2 - 23.5 Seconds
	Turn Rate = $4.5^\circ/\text{Sec}$	34.3 - 27.1 Seconds	16.7 - 14.0 Seconds

Table 2. Time From Intruder Leaving Own Approach Path to the Crossing of the Subject's Approach Path (Intruder Speed 140-180 Knots)

The intruder is initially started at the correct altitude for glideslope intercept, and continues to follow the glideslope until it breaks off its approach (or, in the Missed Intercept case, passes through its approach course) and deviates towards the subject. The vertical rate of the intruder during this blunder is also selected in random order to cover each of three different values during each set of four runs. In one case, the intruder continues his approach descent; in another, the intruder uses a $0.5g$ pull-up to level flight; and in the final case the intruder uses a $0.5g$ pull-up to a climb of 2000 feet per minute.

2.1.4 Test Matrix

The test matrix is shown in Figure 16. Altogether, most subjects had 40 experiment runs, allowing for within-subject comparisons; four subjects did not have runs with the smooth predictor display due to technical problems. The scenarios were flown in 10 blocks of four, where each included all the runs for each particular display-workload combination.

				High Workload	Low Workload
Baseline Display				4 Scenarios	4 Scenarios
Baseline Display	+ Approach Path for Other Runway			4 Scenarios	4 Scenarios
Baseline Display	+ Approach Path for Other Runway	+ Other Aircraft Heading		4 Scenarios	4 Scenarios
Baseline Display	+ Approach Path for Other Runway	+ Other Aircraft Heading	+ 15 Sec. Projection (Noisy)	4 Scenarios	4 Scenarios
Baseline Display	+ Approach Path for Other Runway	+ Other Aircraft Heading	+ 15 Sec. Projection (Smooth)	4 Scenarios	4 Scenarios

Figure 16. Experiment Test Matrix

2.2 Experimental Procedure

The simulator runs with each subject lasted one hour, including briefing, practice runs, all experiment runs, and a debriefing. The briefing explained the displays, controls and procedures involved in the experiment. Subjects were allowed as many practice runs as they requested, and additional practice runs were given before the first experiment runs with each new display. The debriefing questionnaire consisted of simple questions about the subject's background with video games and flying, and of subjective questions about the display attributes; this questionnaire is shown in Appendix A.

The experiment runs each consisted of three sequential parts:

The Flight The subject was told they were flying an approach. Their primary task was to keep their wings level despite turbulence, referenced to an artificial horizon, through the use of a sidestick. The sidestick commands did not affect the path of the aircraft, however, so that consistent approach paths would be followed by the subjects. Their secondary task was to press a red button on the sidestick as soon as they thought the aircraft on a parallel approach was blundering towards them, as evidenced by the traffic display.

The Maneuver Selection Once the subject indicated the parallel approach traffic was deviating towards them, the traffic display was blanked and six possible maneuvers were

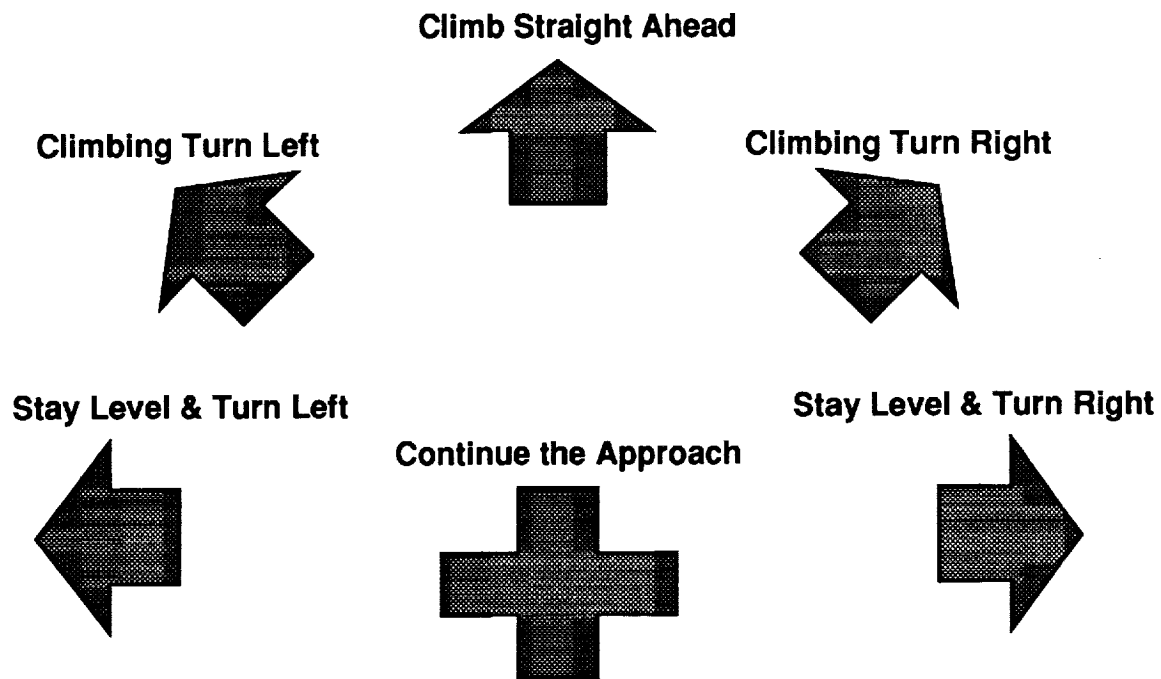


Figure 17. Maneuver Selection Screen

shown to the subject, as shown in Figure 17. Using a mouse, the subject was asked to select the maneuver they considered best for maintaining inter-aircraft separation.

The turn component of each maneuver (when required) used a standard rate turn (3° per second) to a heading 30° off approach course, with a bank rate set such that the heading second derivative was 10° per second squared. The climb component used a 0.5 g pull up to a 2000 fpm climb attitude.

Numerical Simulation of All Avoidance Maneuvers The simulator then calculated the future trajectory of the intruder, and the future trajectory of the subject was calculated for all six possible avoidance maneuvers. Performance metrics of each avoidance maneuver, such as the resulting miss distance, were calculated and stored, simplifying the data reduction process. These numerical simulations were transparent to the subject and did not provide any feedback to them of their performance.

While this does not provide an exact replication of the miss distance achieved by pilots manually controlling the aircraft, it does provide a first order measurement of the subjects' decision making. In addition, regardless of the method of generating the avoidance maneuver, the frequency of collisions found by this experiment can not be considered an estimator of the frequency of incidents expected in real operations, as the intruder trajectories in this experiment may not represent a real intrusion, the defining characteristics of which are not known.

2.3 Simulator Setup

The simulator used a Silicon Graphics Indigo 2 workstation for the displays and aircraft dynamics computations. A sidestick was connected for the flying task, and a mouse for the avoidance maneuver selection. The simulation was designed such that the subject could

easily control their progress, selecting further practice or commencement of the experiment runs.

The aircraft dynamics used simple point-mass calculations with performance constraints representative of air transport aircraft. The pitch steering and heading acquisition models used critically damped controllers, while the localizer acquisition controllers were slightly overdamped, modeling the actual wavering about the approach path of the aircraft.

2.4 Subjects

In total, nineteen subjects flew the experiment. Four subjects tested the first four displays, testing one-quarter of their test matrix; the remaining fifteen subjects tested all five displays, testing the remaining three-quarters of the complete test matrix.

The basic characteristics of the subjects varied widely. Two were current airline flight crew, four were current Certified Flight Instructors (CFI) in general aviation aircraft (one with jet fighter experience), two held Private Pilot Licenses, and the remaining eleven were undergraduate or graduate students.

2.5 Measurements

During each experiment run, several variables were calculated and stored automatically. These include:

- Type of Scenario and Variables Defining Collision Trajectory

- RMS Bank Error from Workload Task

- Time and Aircraft States when Subject Indicates Intruder is Deviating

- Avoidance Maneuver Selected by Subject, and Time Required to Select It

- Achieved Performance of All Avoidance Maneuvers, Including Minimum Miss Distances (Slant Range, Horizontal and Vertical).

In addition, subjective comments were elicited once all the scenario runs were complete using the Questionnaire shown in Appendix A.

3. Experiment Results

The results of this experiment will be discussed as follows. First, the overall performance of the subjects will be discussed. Then the comparative effects of each of the elements in the test matrix (displays, workload and scenarios) will be discussed and an analysis of the effects of different collision geometries will be given. Finally, the metrics of subjects with different characteristics will be discussed.

3.1 Overall Performance

Several measures exist for examining the validity and speed of the subjects' determination that the intruder was deviating towards their own approach path. First, as shown in Table 3, the correctness of the subjects' decisions on when to generate an alert can be evaluated. Table 3 is divided into two parts, one showing the correctness of the decisions when the traffic blundered off its approach path, the other showing the decisions when the traffic did not blunder but instead maintained its approach path. A near-miss/collision is defined as a miss distance less than 500 feet, the same metric used in evaluation of other collision avoidance systems for parallel approaches [Pritchett et al, 1995; Ebrahimi, 1993]

In the cases where the intruder was scripted to deviate, the subject correctly spotted the anomaly in time to safely avoid the other aircraft almost 82% of the time. However, some deviations were not noticed by the subjects until the aircraft spacing was less than 500 feet (0.7% of the runs), and others were noticed so late that any avoidance maneuver would result in a near-miss/collision (16.4%). Subjects also generated an alert, in some runs, while the intruder was maintaining his proper approach course (1.0% of the runs). This classification of reactions only evaluates the timing of the alert, and does not consider the possible effectiveness of an ensuing avoidance maneuver.

In the cases where the intruder was not scripted to deviate, the subjects' correctly did not generate an alert over 97% of the time. False alarms were given in 2.9% of the cases.

Type of Reaction	Description	% of Approaches
Scenarios Involving Intruder Deviation Towards Subject		
Correct Detection	A potential collision is spotted in advance.	81.9%
Missed Detection	A near-miss/collision is spotted only after it has occurred	0.7%
Late Alert	An alert is given before a near-miss/collision, but too late safely avoid the other aircraft	16.4%
Early Alert	The subject alerts while the intruder is still maintaining their correct approach course	1.0%
Scenarios Without an Intruder Deviation		
Correct Rejection	Subject correctly does not generate an alert	97.1%
False Alarm	The subject alerts although the intruder will maintain their correct approach course	2.9%

Table 3. Frequency and Types of Alerting Decisions

Next, the average reaction time of the subjects, from when the intruder started to leave its approach path to when they generated an alert, ranged from -1.96 seconds (for a case when the intruder was missing its intercept and would soon pass through its approach path) to 31.69 seconds, with a mean of 9.73 seconds. With these reactions, the estimated time left until the point of closest approach ranged from -13.39 (the subject reacted after the point of closest approach) to 34.32, with an average of 14.37 seconds

Like the preliminary experiment, no apparent alerting logic can be found, other than an appearance of alerting based on range or lateral separation. Figure 18 shows a scatter plot of the range and range rate for all the subject reactions; the data appears to be widely scattered, has a low correlation and does not appear to follow a TCAS-like algorithm based on predicted time to collision. Figure 19 shows a histogram of the lateral separation of the aircraft when an alert was generated; subjects appeared to alert when the lateral separation between the aircraft was between 1000' and 2000' feet, with an average of 1340 feet.

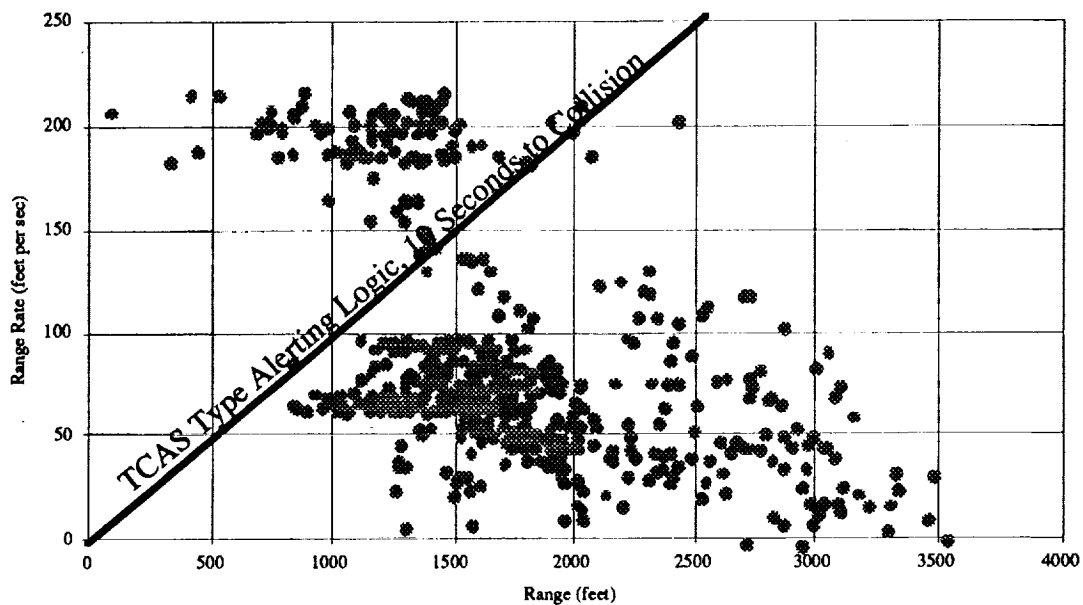


Figure 18. Range versus Range Rate Plot When Subjects Generated an Alert

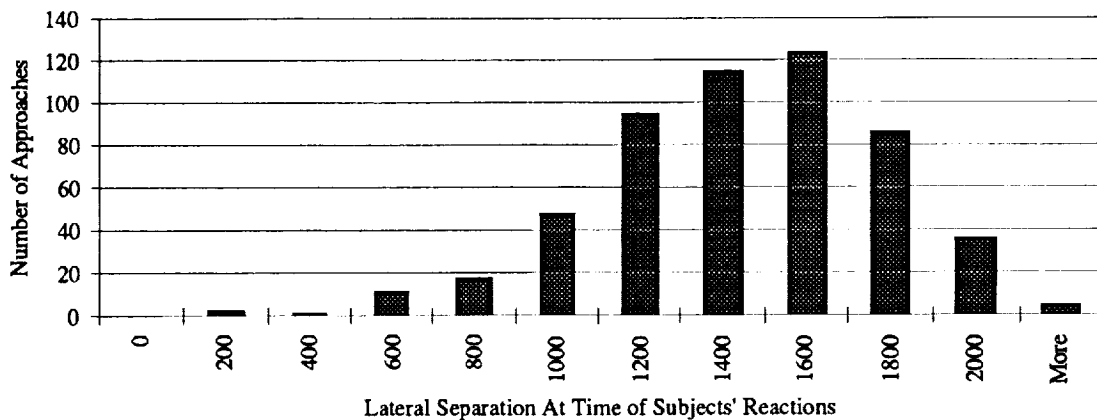


Figure 19. Histogram of Aircraft Lateral Separation When Subjects Generated Alert

In addition to the timing and validity of the subject's alerting decisions, the performance of the subjects in selecting a safe direction of flight for an avoidance maneuver can be evaluated. The subjects were asked to select one of six possible avoidance maneuvers, and the performance of all six were calculated numerically.

Table 4 lists the frequency with which subjects selected each avoidance. The most popular maneuvers were Turn Away and Climb, and Turn Away (Maintaining Altitude), showing a strong preference for turning-away maneuvers. The remaining maneuvers were selected rarely.

Maneuver	Frequency Maneuver Was Selected
Turn Away from Intruder (Maintain Altitude)	36%
Turn Away from Intruder and Climb	55%
Climb (No Turns)	3%
Turn Towards Intruder and Climb	2%
Turn Towards Intruder (Maintain Altitude)	2%
Continue the Approach	2%

Table 4. Frequency of Avoidance Maneuvers Selection

The effectiveness of each avoidance maneuver during all runs can be compared with its effectiveness in the specific runs when selected by the subject. The subjects selections caused near-misses/collisions in 23% of the experiment runs. If the subjects had not reacted but instead had continued on their approach, near-misses/collisions would have resulted in 43% of the approaches, and if they had chosen maneuvers randomly, near-misses/collisions would have resulted in 38% of the approaches. Therefore, the subjects' selections caused a significant improvement in collision avoidance, although the final collision rate was still high.

Finally, the average bank deviation found with the subjects wings-leveling task can be analyzed. Specifically, the average bank deviation after the blunder commenced (4.39°) was found to be very significantly higher ($p < 0.01$) than the bank deviation averaged over the entire run (3.68°), suggesting that the subjects partially abated their wings-leveling task in an attempt to discern if an alert was required.

3.2 Display Effects

In this experiment, five different displays were tested, as listed in Table 5. Some display effects were found, and will be described using the same metrics as for the overall results.

Display Name	Display Description
Baseline Display	Based on TCAS II Type Display on Moving Map
Feducial Mark Display	Adds Indication of Other Aircraft's Approach Path Location
Heading Display	Adds Indication of Other Aircraft's Heading
Noisy Projection Display	Adds 15 Sec. Projection of Other Aircraft's Position Based on Noisy Measurements
Smooth Projection Display	Adds 15 Sec. Projection of Other Aircraft's Position Based on Exact Measurements

Table 5. Displays Tested in the Experiment

First, the correctness of the subjects' alerting decisions is summarized in Table 6. All of the false alarms generated during the non-blunder scenarios occurred with the Baseline Display and with the Noisy Projection Display. All of the early alerts generated before a blunder started occurred with the Noisy Projection Display, a significant difference ($p < 0.05$). Runs with both the Baseline and Feducial Mark displays resulted in one missed detection, while the runs with the Noisy Projection display resulted in two missed detections, a difference that is not statistically significant.

Display Type	Intruder Deviates Toward Subject				No Intruder Deviation	
	Correct Detection	Missed Detection	Late Alert	Early Alert	Correct Rejection	False Alarm
Baseline	82.3%	0.9%	16.8%	0%	94.4%	5.6%
Feducial Mark	83.3%	0.9%	15.8%	0%	100%	0%
Heading	83.3%	0%	16.7%	0%	100%	0%
Noisy Projection	77.2%	1.8%	15.8%	4.4%	91.7%	8.3%
Smooth Projection	84.3%	0%	15.7%	0%	100%	0%

Table 6. Correctness of Subject Alert Generation

As shown in Figure 20, a significant reduction in reaction time was found with all the newer displays compared to the Baseline display ($p < 0.05$). A corresponding increase in the time remaining after the alert until the collision was also found. By both metrics, the Noisy Projection display performed the best, at the expense of its higher false alarm rate.

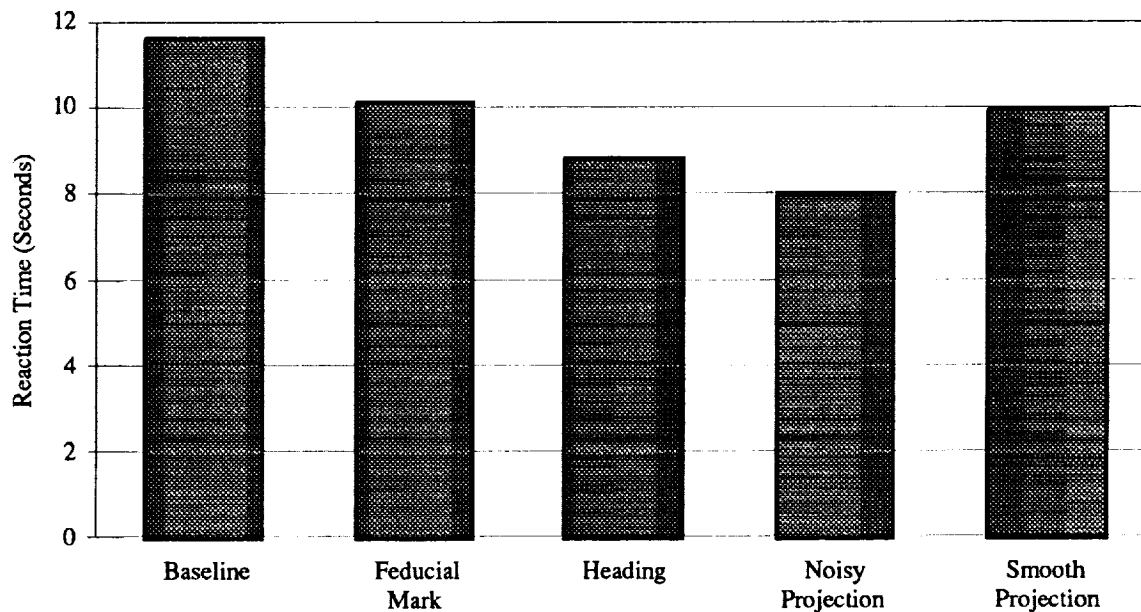


Figure 20. Average Reaction Time With Each Display

Although the newer displays were purposefully designed to give indications of relative convergence rate and trend, no differences can be found in the method used by the subjects to generate alerts with each of the different displays. Very little correlation between the inter-aircraft range and the convergence rate can be found at the time of the subjects' alerts, showing subjects did not tend to alert when a specific time to collision was remaining.

Instead, the subjects appeared to use an abnormal lateral position of the intruder aircraft from its approach position as a criteria for generating alerts.

Although the subjects appeared to use this same method for all displays, a difference can be found in that the lateral deviation of the intruder aircraft from its own approach path required to generate an alert was very significantly smaller ($p < 0.01$) with the newer displays compared to the baseline display. This effect is shown in Figure 21, and it correlates to the quicker reaction times with these displays.

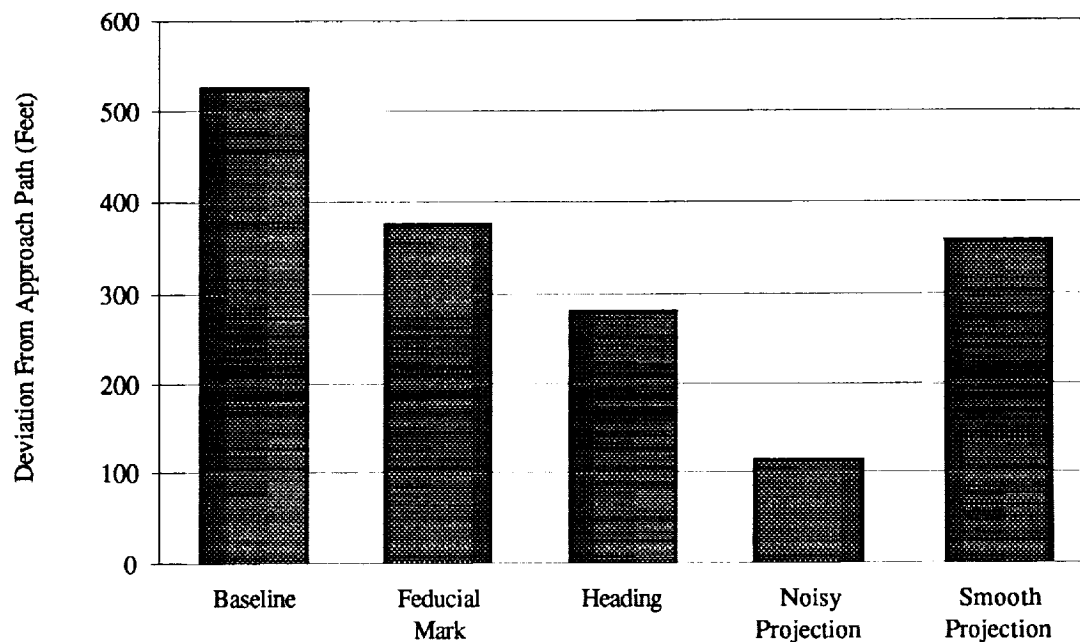


Figure 21. Average Intruder Deviation From Feducial Mark at Subject Alert Times With Each Display

Few significant differences can be found in the avoidance maneuvers selected by the subjects. Each maneuver appeared to be selected the same amount, regardless of display. As shown in Figure 22, the frequency of near-misses/collisions appears to be reduced with the newer displays; however, none of these differences are statistically significant.

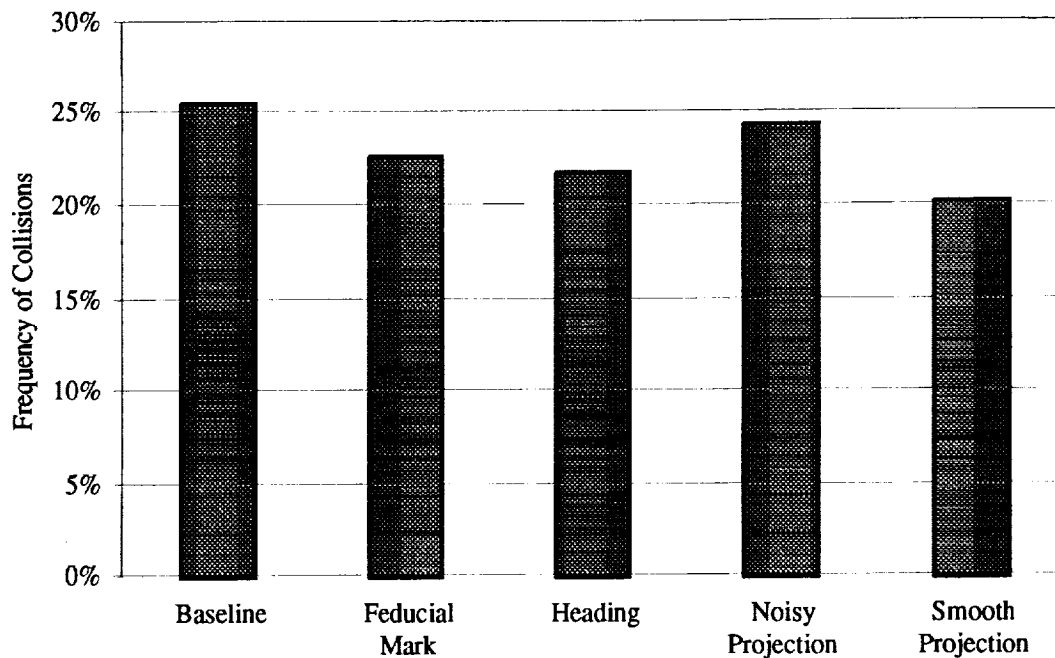


Figure 22. Frequency of Near-Misses / Collisions with Each Display

3.3 Workload Effects

Concurrent with monitoring a traffic display for possible traffic incursions, subjects were responsible for a wings-leveling task, using an artificial horizon and a side-stick. The difficulty of this task was controlled by generating high or low amounts of turbulence in bank, and thereby creating a high or low workload for the subject to attend to away from the traffic display.

Most of the performance measures for this experiment were nearly identical when comparing the data from runs with high workload against runs with low workload. For example, the average reaction time in the high workload runs was 9.74 seconds, compared with 9.73 seconds for the low workload runs.

Only two metrics showed noticeable variation between the high and low workload cases. First, as shown in Figure 23, the frequency of near-misses / collisions in the high workload was 26%, compared to 20% in the low workload cases. This difference, however, can only be tested significant to the 80% level.

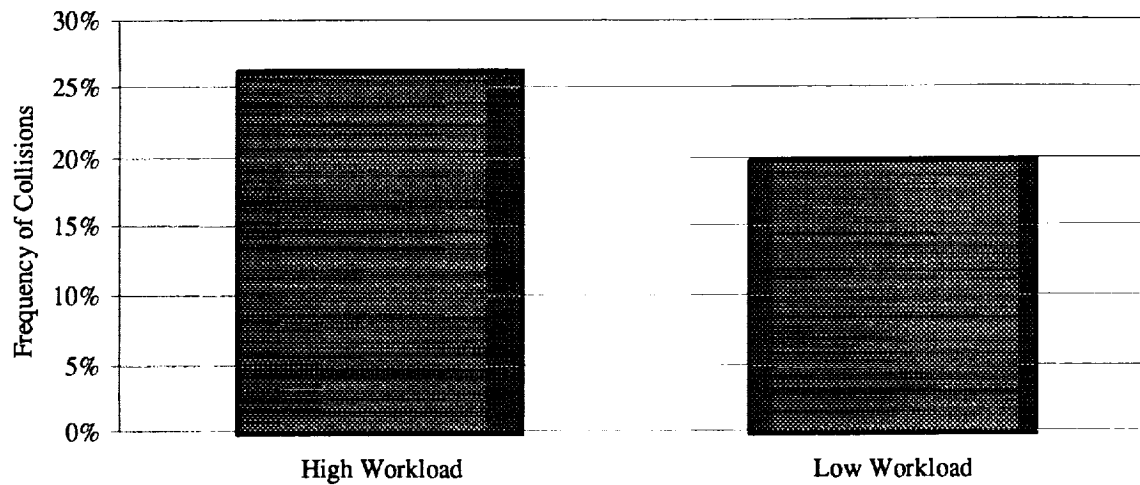


Figure 23. Frequency of Near-Misses / Collisions with Different Subject Workloads

Second, the mean squared error in the wings-leveling task itself was very different, as shown in Figure 24, especially for the period of time after the blunder had started. This illustrates the comparative difficulty of the different workloads.

The large difference between overall MSE bank and MSE bank once the blunder had started may indicate that, in high workload conditions, the subjects decided to drop the wings-leveling task in order to adequately assess the traffic situation. This differs from the subjects' briefings, in which they were asked to consider the wings-leveling task to be primary.

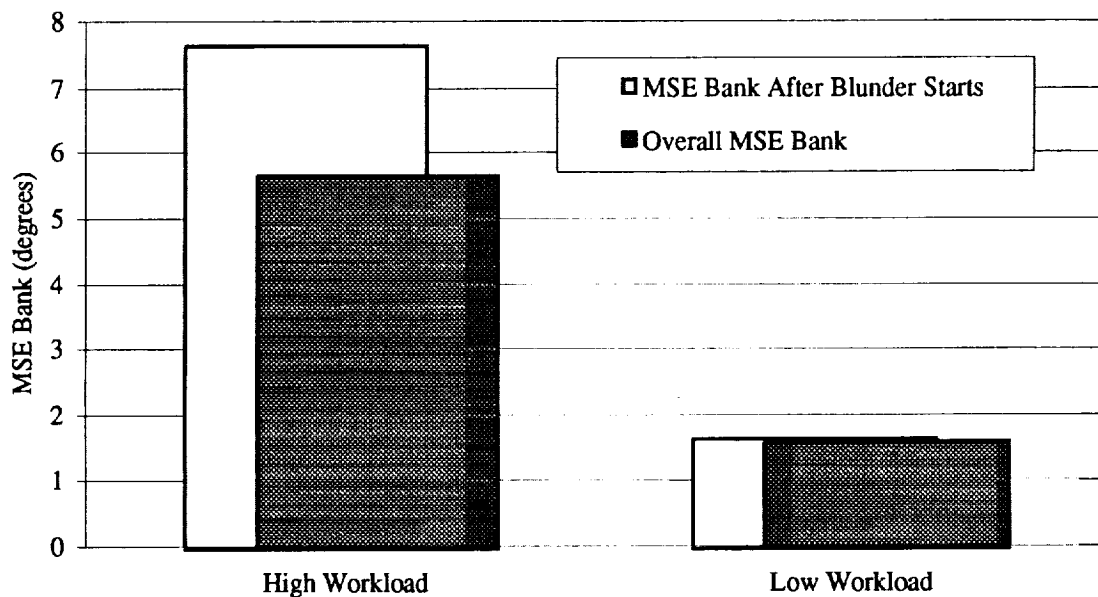


Figure 24. MSE Bank for Different Subject Workloads

3.4 Collision Geometry Effects

The final independent variable in this experiment's test matrix was to vary the intruder trajectories. Within each set of four runs, the intruder trajectories were one each of: Safe Approach, Missed Intercept, Hazardous Blunder and Less-Hazardous Blunder.

Within the three types of scenarios involving a deviation of the intruder towards the subject, the commanded state variables defining the intruder's trajectory were also controlled. The convergence rate was varied between high and low values (45° and 15° intercept angles respectively), and the intruder's vertical speed was varied between a continued descent, level flight, and a 2000 fpm climb.

Within the two Blunder scenarios, where the rate of turn was also a factor, the heading rate was varied between high and low values ($4.5^\circ/\text{second}$ & $1.5^\circ/\text{second}$ respectively).

3.4.1 Differences in Performance Between the Different Scenarios

The average reaction time to the Missed Intercept scenarios was significantly lower than that for both type of Blunder scenarios ($p < 0.01$), as shown in Figure 25. This may be partially an artifact of the way reaction time is defined, however; for the Missed Intercept scenarios, the measurement is from the instant the intruder crosses his approach path with a fully established intercept angle, whereas the Blunder scenario measurement starts when the intruder starts to turn away from its approach heading. Even with the lower reaction times, the subjects' reactions in the Missed Intercept scenarios tended to leave them with a significantly lower expected time to collision than in the Blunder scenarios because of the shorter duration of the intruder's collision course, for this type of scenario, once it has passed inside its own approach course.

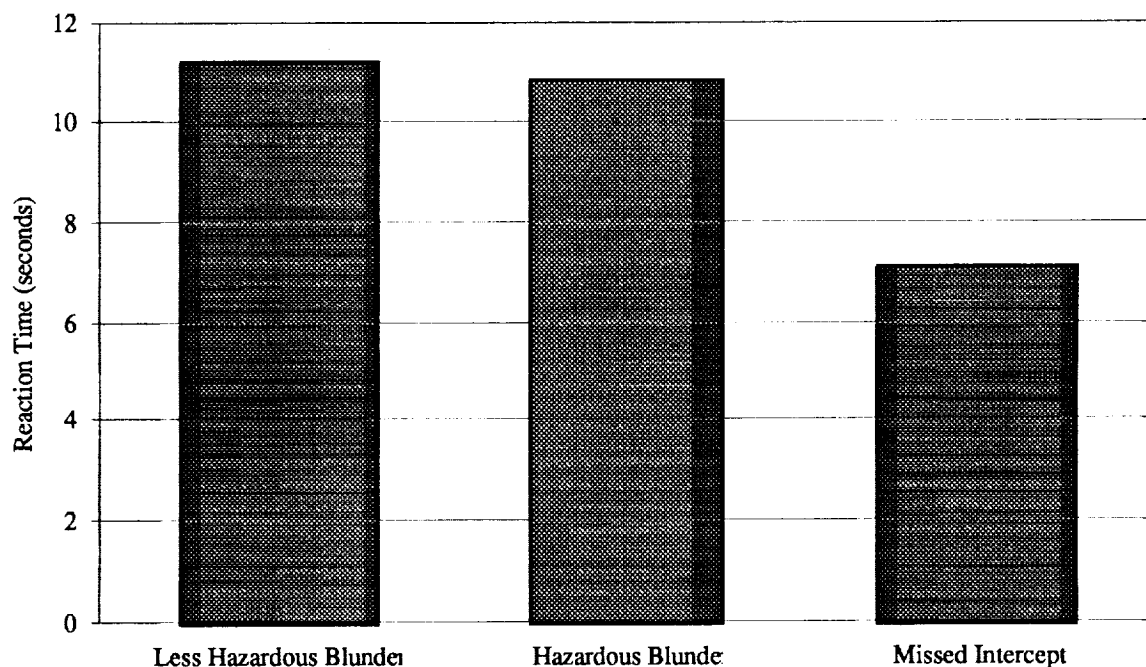


Figure 25. Average Reaction Time with the Different Scenario Types

Congruent with their earlier reaction times, the Missed Intercept scenarios had subject-generated alerts when the intruder aircraft had deviated less from its approach path, as shown in Figure 26. All differences are significant ($p < 0.05$).

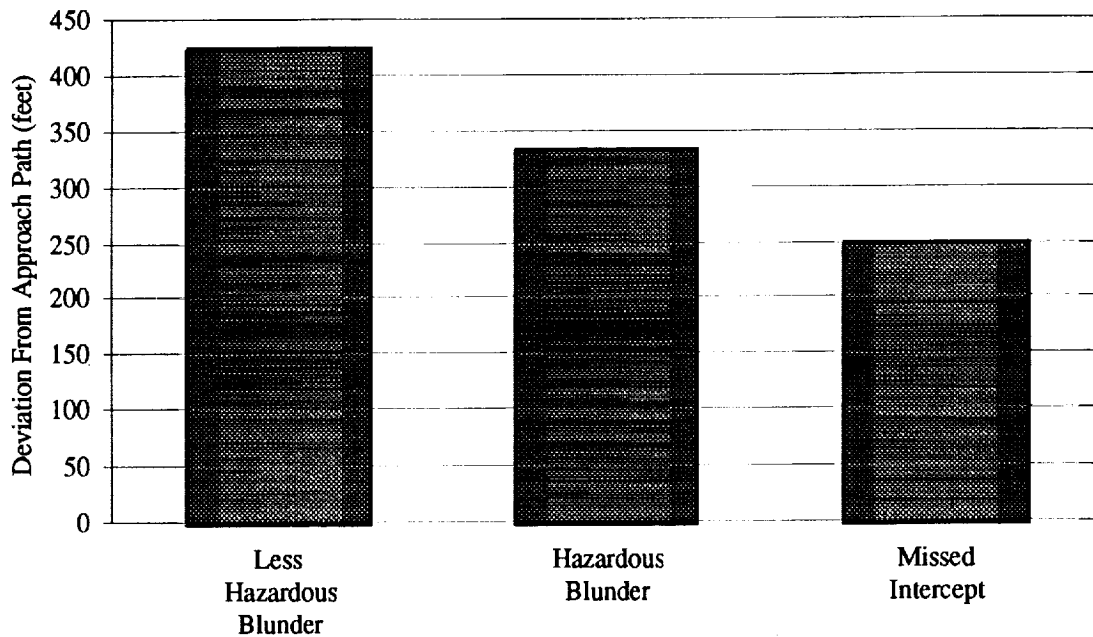


Figure 26. Average Intruder Deviation From Approach Path with the Different Scenario Types

The avoidance maneuvers were each selected with the same frequency across all scenarios, showing that the subjects did not apply a different algorithm for the different intrusion geometries. However, as shown in Figure 27, the subjects achieved significantly different performance in avoiding near-misses / collisions in the different scenarios ($p < 0.01$). No collisions were found for the Less-Hazardous Blunder scenarios; the intruder was sufficiently far away from the subject in these scenarios that any avoidance maneuver would be safe. Near-misses / collisions occurred in 19% of the Hazardous Blunders, an improvement from the 51% that would have occurred if the subject had done nothing. Finally, near-misses / collisions occurred in 51% of the Missed Intercept scenarios, an improvement from the 78% that would have happened if the subject had done nothing.

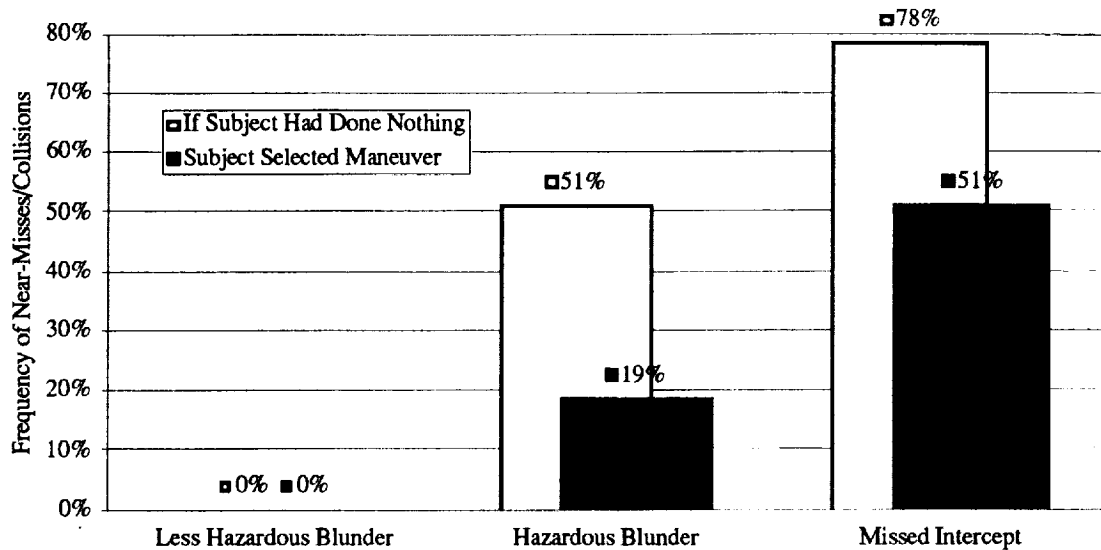


Figure 27. Frequency of Near-Misses / Collisions with the Different Scenario Types

3.4.2 Differences in Performance Based on Aircraft Convergence Rate

For all of the three types of scenarios involving an intruder deviation towards the subject, the convergence rate between the aircraft was controlled by setting the intruder intercept heading angle. In one half of the experiment runs the intercept angle was set to be high (45°) and in the remaining one half the intercept angle was set to be low (15°).

An intrusion with a high convergence rate created a much more time critical situation than an intrusion with a low convergence rate. This was shown by the average duration of the intrusion, measured from when the intruder first left its approach path until when the intruder crossed the subjects approach path; the average blunder duration for a high convergence intrusion was 16.84 seconds, about half of the 31.15 second duration for a low convergence intrusion.

The correctness of the subjects' alerts are shown in Table 7. All of the missed detections and late alerts occurred with the high convergence intrusions. An approximately equal number of early alerts happened with each, as would be expected.

	Correct Detection	Missed Detection	Late Alert	Early Alert
High Convergence Rate	66.9%	1.5%	32.4%	0.7%
Low Convergence Rate	98.9%	0%	0%	1.1%

Table 7. Correctness of the Subjects' Alerts to the High and Low Convergence Intrusions

The subjects' tended to be much quicker to notice a high convergence rate intrusion, as shown in the histograms of reaction times in Figure 28. The average reaction time to a high convergence intrusion was 8.44 seconds, which is significantly lower ($p < 0.01$) than 11.02 seconds, the average reaction time to a low convergence intrusion.

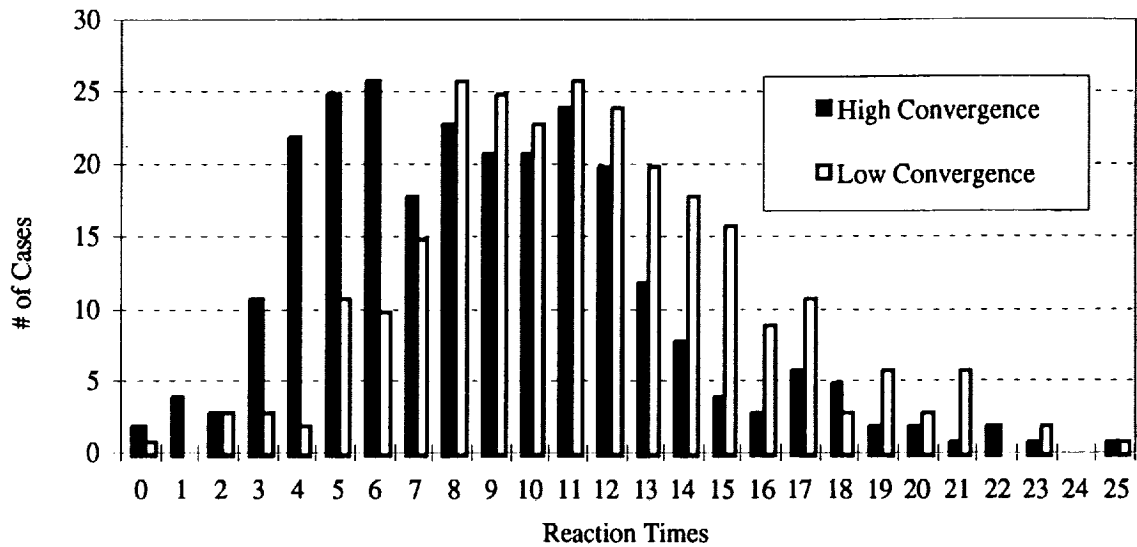


Figure 28. Histograms of Reaction Time to High and Low Convergence Intrusions

Analysis of the criteria the subjects used to generate an alert found that the same lateral separation between the aircraft existed at the subjects' reaction times. The histograms of the lateral separation at the subject's alert are shown in Figure 29. There is no statistically significant difference between them.

This suggests the pilots used lateral separation as the primary alerting criteria; the earlier reaction times with the high convergence intrusions would be caused by the intruder reaching the critical lateral position earlier.

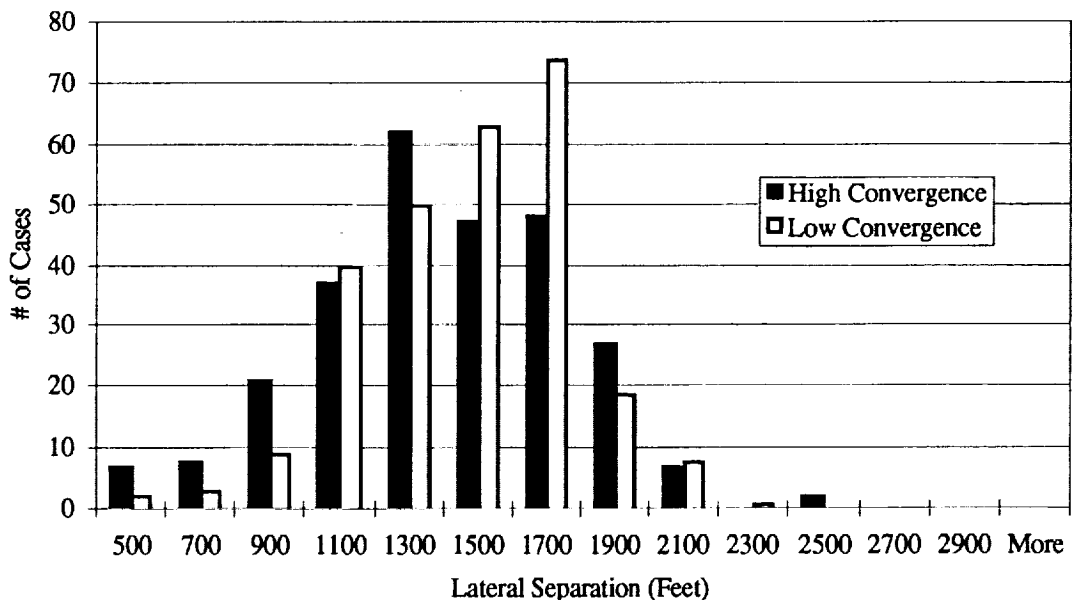


Figure 29. Histograms of Aircraft Lateral Separation at Subjects' Reactions

As shown in Figure 30, the subjects tended to select different avoidance maneuvers based on the intruder convergence rate. In high convergence rate situations, the subject was more likely to choose Turn Away and Climb ($p < 0.01$) or Turn Into and Climb ($p < 0.05$). In low convergence rate situations, Turn Away was more likely ($p < 0.01$).

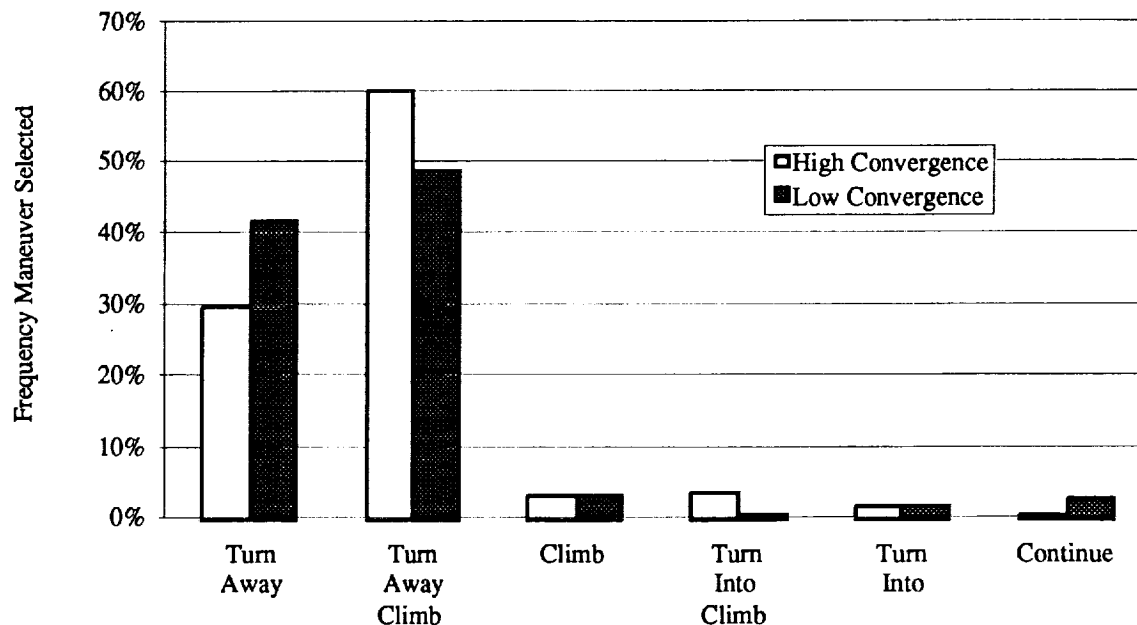


Figure 30. Frequency of Avoidance Maneuver Selection with Inter-Aircraft Convergence Rate

The ensuing frequency of near-misses / collisions differed greatly between the high and low convergence rate runs. As shown in Figure 30, the near-miss / collision rate for the high convergence rate runs was 42%, which is a higher near-miss / collision rate than would be found if the subjects had always continued the approach (39%). The near-miss / collision rate for the low convergence runs, 4%, is significantly lower than that for the high convergence runs ($p < 0.01$); this rate is also significantly better than that achieved by always continuing the approach (47%).

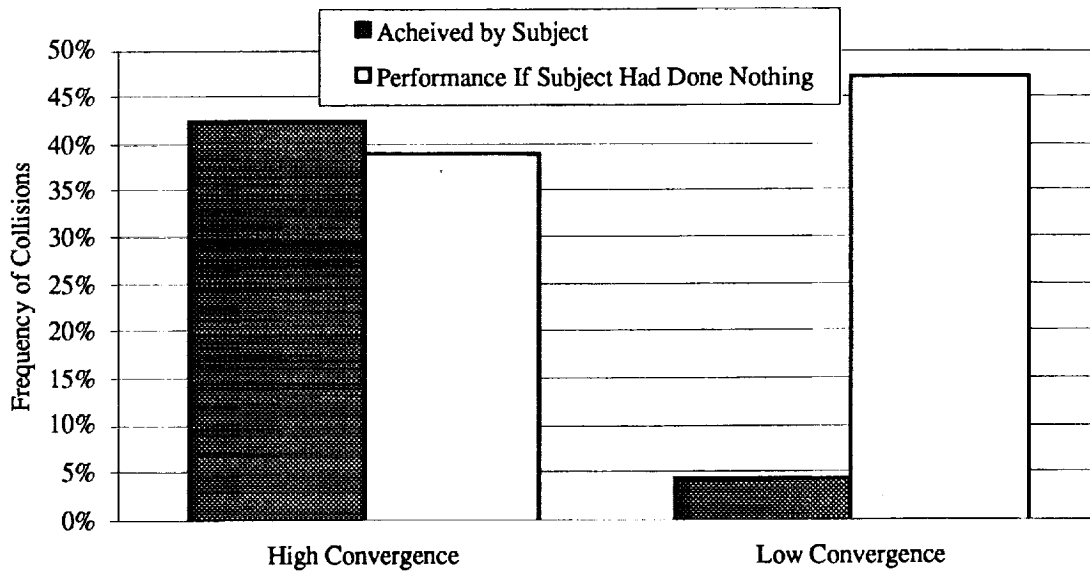


Figure 31. Frequency of Collisions with High and Low Convergence Intrusions

3.4.3 Differences in Performance Based on Turn Rate

Both the Hazardous Blunder and Less Hazardous Blunder type scenarios controlled the turn rate the intruder used to establish its intercept angle, setting it to a high turn rate (4.5° per second) once half the time, and to a low turn rate (1.5° per second) the other half. These two values, combined with the two values used for the intercept angles, provided four possible heading and turn rate combinations. Each of these four combinations were used once in the four total blunder scenarios given with each display.

As shown below in Table 8, each combination of Turn Rate and Convergence Rate caused a dramatically different total duration of a blunder, measured from when the intruder first turned off its approach course to when it reached the subject's approach course. These values created different timescales in the subjects' task of recognizing a blunder, and also appeared as very different pictures on the traffic displays.

Convergence Rate & Turn Rate	Average Total Blunder Duration
High Convergence Rate, High Turn Rate	15.33 Seconds
High Convergence Rate, Low Turn Rate	24.69 Seconds
Low Convergence Rate, High Turn Rate	30.54 Seconds
Low Convergence Rate, Low Turn Rate	33.70 Seconds

Table 8. Total Blunder Duration for Different Convergence & Turn Rates

The reaction times tended to become significantly better for those runs with the higher convergence and turn rates, as shown in Figure 32. Those with the high convergence and turn rate had an average reaction time significantly lower than all the others ($p < 0.01$); those runs with the low convergence and turn rates had an average reaction time significantly higher than all but the high convergence, low turn rate cases ($p < 0.01$).

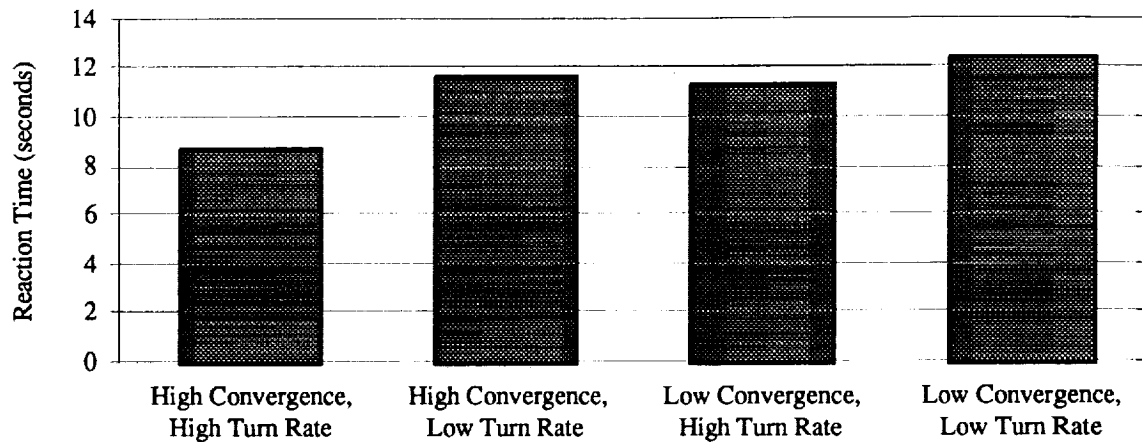


Figure 32. Average Reaction Times with Different Convergence and Turn Rates

As was found in section 3.4.2, the higher reaction times may indicate the subjects' alerting criteria is based on a relatively constant lateral separation. Figure 33 shows the histograms of the intruder lateral deviation from its approach path when at the point the subjects generated an alert. There is no statistically significant difference between the distributions for the different rates.

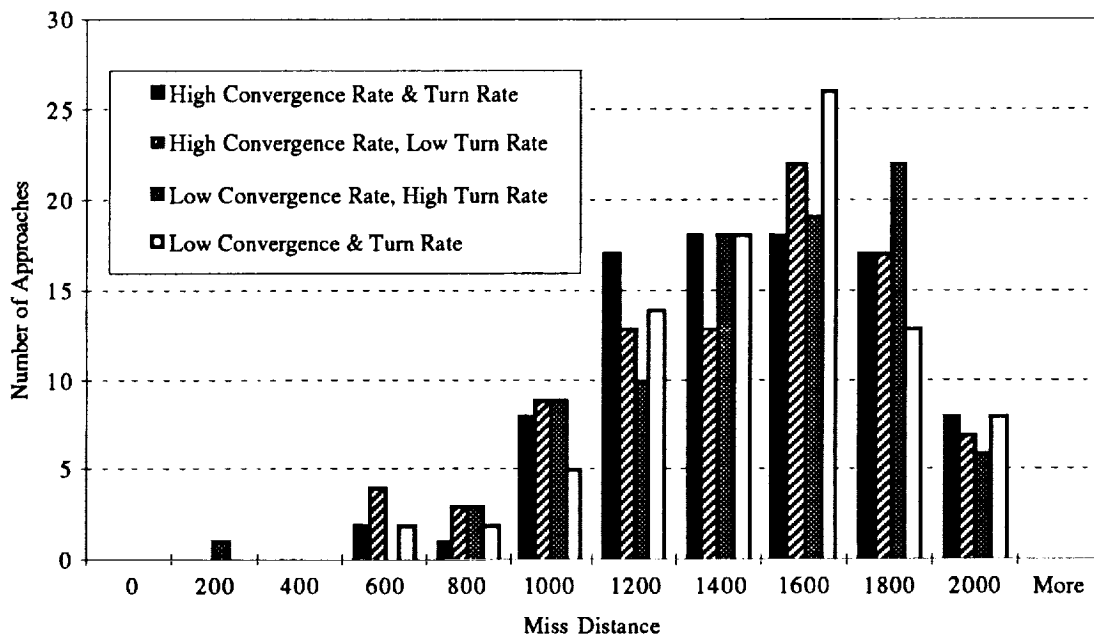


Figure 33. Histogram of Intruder Lateral Deviation From Nominal Approach Path at Subject Reactions with Different Convergence and Turn Rates

The difference between the subjects' alerting logic and the type of logic used by current TCAS-type systems is illustrated in Figure 34, a scatter plot across all the Blunder scenarios comparing reaction time to time remaining before collision. A TCAS-type alerting logic generates an alert a fixed amount of time before an anticipated collision, which can be represented on this scatter plot as a vertical line at the appropriate time to collision. The subjects had a wide amount of scatter along all the possible reaction times.

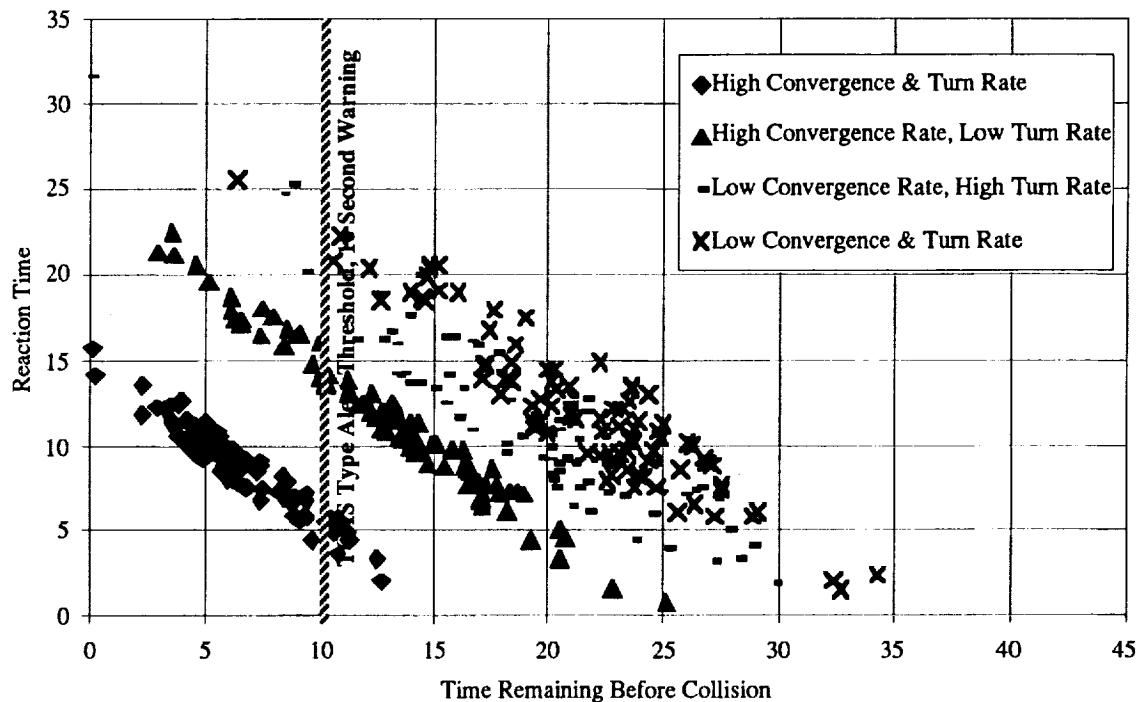


Figure 34. Scatter Plot Comparing Reaction Time and Time Remaining with Different Convergence and Turn Rates

Finally, significant differences in the frequency of near-misses / collisions were found based on convergence and turn rate, as shown in Figure 35. The runs with the high convergence and turn rates incurred the worst performance, a difference compared to all other cases that is very significant ($p < 0.01$). The other cases are not significantly different amongst themselves.

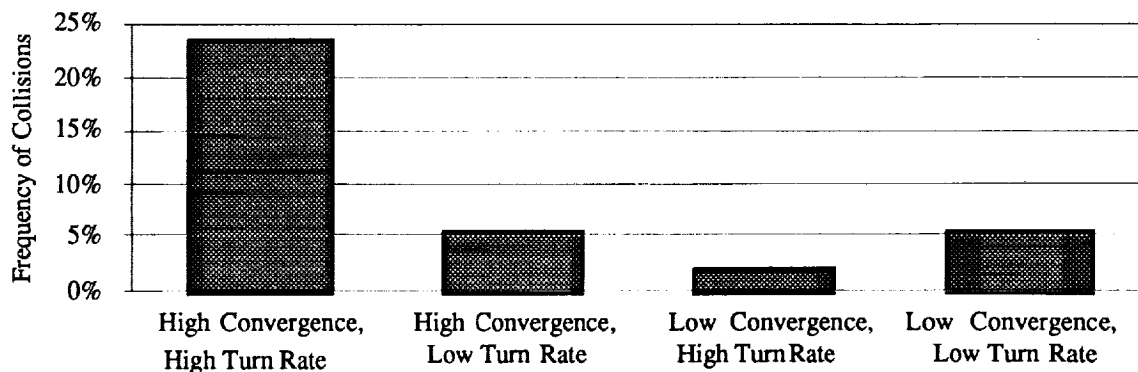


Figure 35. Frequency of Near-Misses / Collisions with Different Convergence and Turn Rates

3.4.4 Differences in Performance Based on Intruder Vertical Speed

The final variation in intruder trajectory was the intruder's vertical speed once it left the approach path. Until the intruder deviated, it maintained an approach descent along its glideslope - it then followed one of three vertical speeds: a continued approach descent, level flight, or a climb at 2000 feet per minute.

Unlike the horizontal path of the intruder, the vertical flight path was not graphically shown to the pilot. The only representation of the intruder's altitude was a small three digit text number next to the intruder's icon, showing whether the intruder was above or below the subject, and the altitude difference in discretized units of hundreds of feet. This representation models the current type of TCAS displays, and is a limitation of a two dimensional traffic representation.

The intruder's vertical speed was not shown, and could only be inferred by the rate of change of the textual altitude display. Because the altitude was discretized to hundreds of feet, the time required for a change in the text indication could be several seconds.

The visual picture of the intruder's vertical speed at the subjects' reaction times were often not very compelling. Because the intruder's new vertical speed required time to establish, the intruder's relative altitude and its absolute vertical speed at the time of the subjects' reaction depended strongly on reaction time and therefore followed its variance. Many runs, especially those with an early reaction, presented a relative altitude and absolute vertical speed similar to the co-altitude condition found during an approach descent, as shown in Figures 36 and 37.

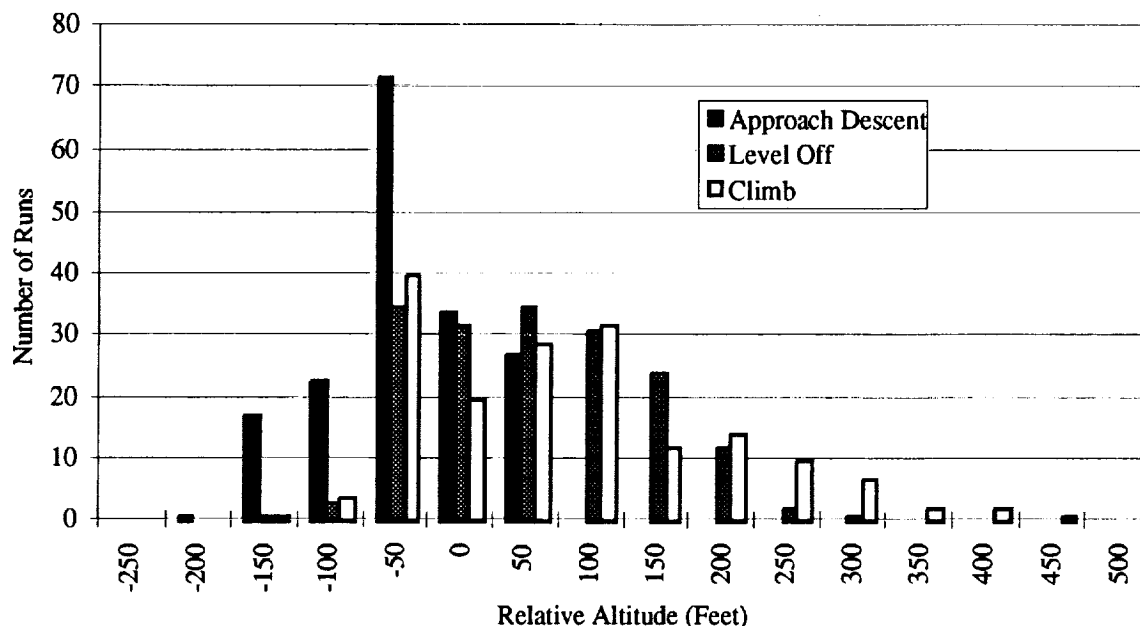


Figure 36. Intruder Relative Altitude at Subject Reaction Time, by Different Commanded Vertical Speeds

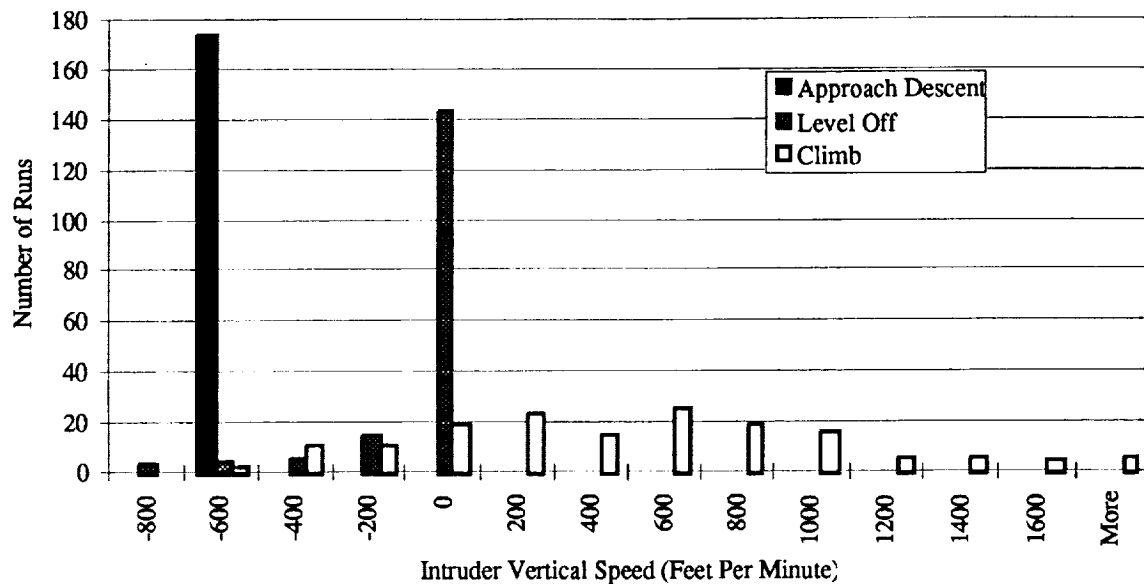


Figure 37. Intruder Vertical Speed at Subject Reaction Time, by Different Commanded Vertical Speeds

All of the performance metrics about the reaction times and correctness did not vary with different intruder commanded vertical speeds. The average reaction times, and the lateral separation where they occurred, can not be tested to be significantly different.

Differences can be found, however, in the amount subjects tended to select different avoidance maneuvers. As shown in Figure 38, the subjects tended to choose the level Turn Away maneuver significantly more often when the intruder was commanded to level off or climb ($p < 0.05$). Conversely, the subjects tended to choose the Turn Away Climb maneuver significantly more often when the intruder was continuing his approach descent. This indicates the subjects were often aware of the intruder's vertical trend. However, dangerous maneuvers, such as the climbing maneuvers when the intruder was climbing, were still frequently selected, showing the subjects' awareness of the intruders' projected vertical flight path was not consistent.

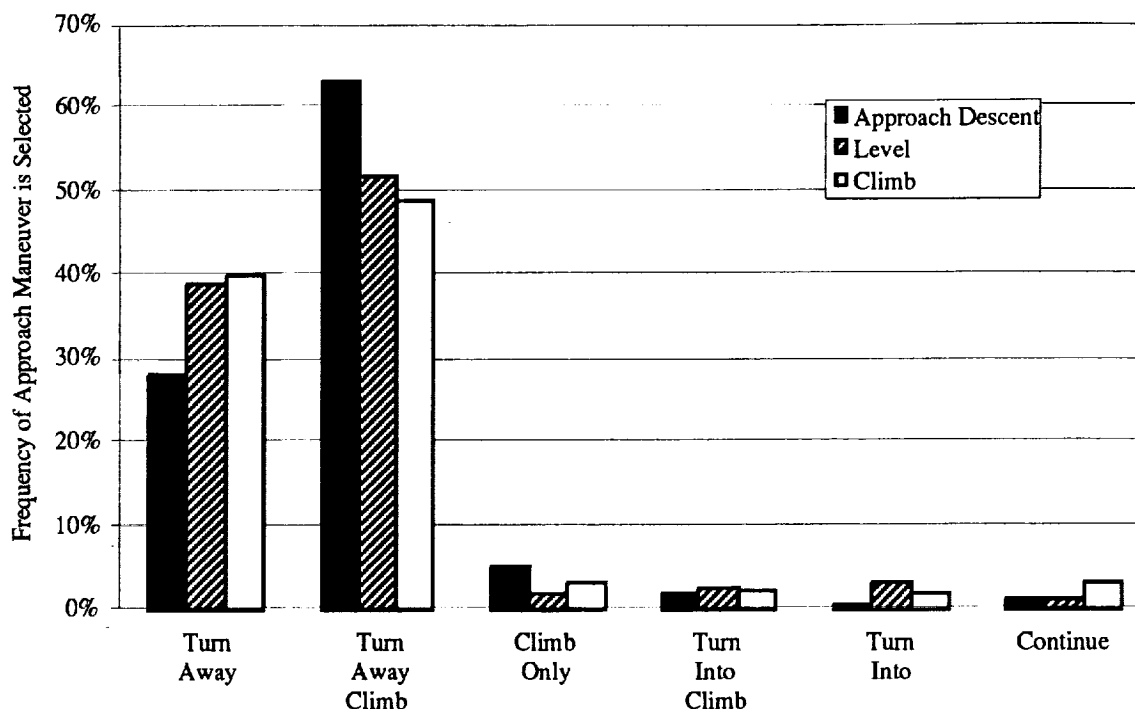


Figure 38. Frequency of Avoidance Maneuver Selection, by Different Commanded Vertical Speeds

No significant differences were found in the near-miss / collision frequency between the runs with different intruder vertical speeds.

3.5 Variance With Subject Characteristics

Differences in performance between subjects of different characteristics were examined. Several measures were compared, including whether or not the subject was a pilot, and whether or not the subject described themselves as a 'Video Game Junkie'.

With these comparisons, no statistically significant differences could be found in any of the standard performance metrics. The similar performance of pilots and non-pilots may suggest that this particular type of conflict detection is not something pilots are currently trained for or accustomed to.

Subjects were also to describe their age in terms of an age group. The results of the following groups were compared: Young (age less than or equal to 25), Older (age greater than 30), and Middle. These age groups are younger than those used to describe the population at large, mostly due to large number of undergraduate and graduate students who flew the experiment.

In general, the younger subjects tended to react significantly quicker ($p < 0.01$), as shown in Figure 39. A corresponding decrease in intruder deviation at reaction time was also found. However, this does not represent an increase in accuracy, as they accounted for 80% of the false alarms, a significant difference ($p < 0.05$).

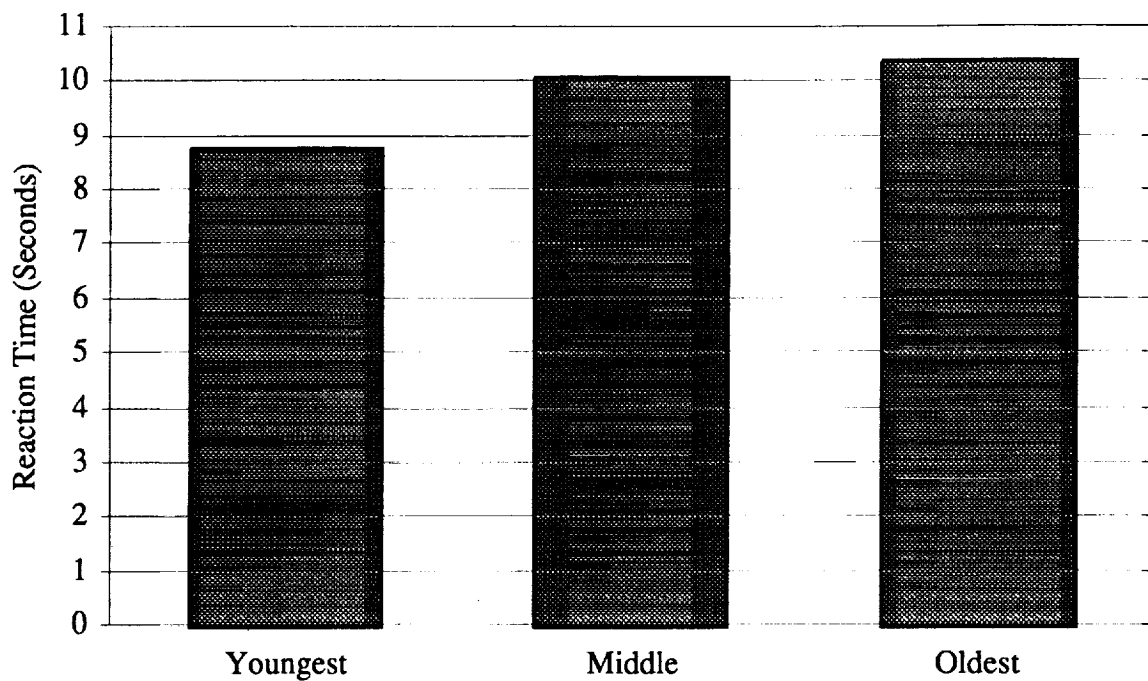


Figure 39. Average Reaction Times of Different Age Groups

The younger subjects were also better at the task of keeping their wings level, as shown in Figure 40. Their performance is significantly better ($p < 0.01$).

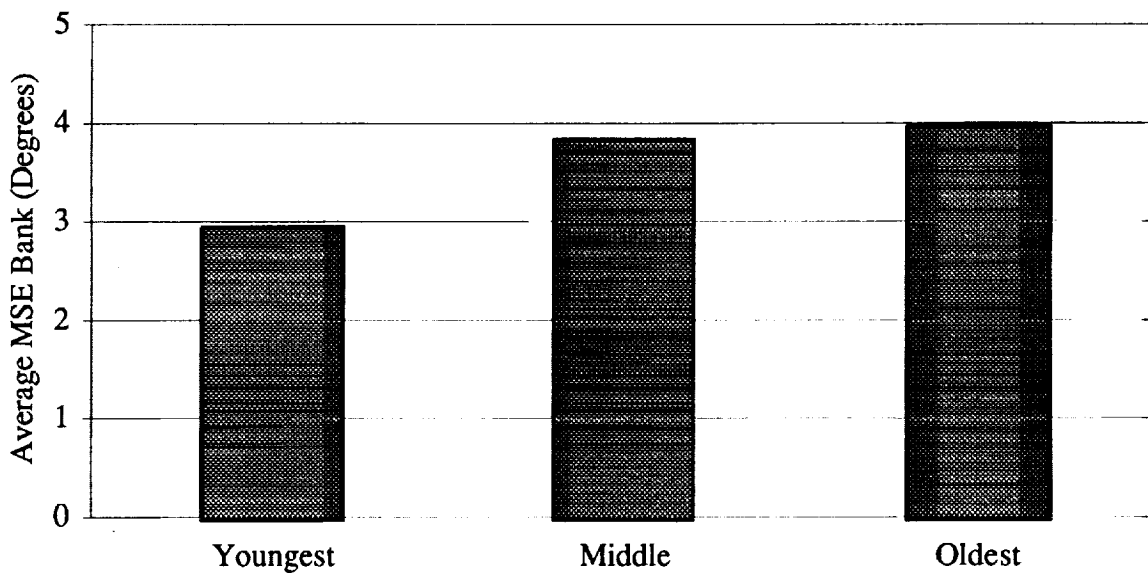


Figure 40. MSE Bank During Wings-Leveling Task of Different Age Groups

4. Conclusions

- Subjects' alerting schema consistently used the lateral deviation of the intruder from its approach path as the primary criteria for generating an alert. The display of heading and trend information did not appear to encourage different, more effective alert logic schemes; instead, the subjects used the same alert logic with lower thresholds.
- Subjects dropped the sidetask once the intruder started deviating, indicating the task of monitoring traffic and deciding when to alert requires a pilot's full attention at critical times.
- Between the displays tested, no significant reductions in near- miss / collision frequency were found, suggesting the newer displays did not help the subjects predict which would be the safest direction of flight for an avoidance maneuver.
- The traffic displays primarily presented the horizontal traffic situation; the vertical trajectory of the intruder could only be inferred through a simple text display of relative altitude. Combined with the co-altitude conditions at the start of the intruder's deviation towards the subject, this made an accurate assessment of the projected vertical trajectory of the intruder difficult. Although subjects appeared to often be aware of the intruder's altitude trend, they more often picked avoidance maneuvers that did not generate vertical separation.
- The similar performance of pilots and non-pilots as subjects in this experiment suggests that recognition of a potential traffic conflict during closely spaced parallel approaches is not a task pilots are accustomed to. Training may improve pilots' performance.
- Selecting a direction of flight for an avoidance maneuver appears to be a complex, time-critical task. Subjects had very little time before a potential collision, and the safest maneuver depended on subtle variations in the intruder trajectory. Subjects tended to pick the maneuvers Turn Away from the Intruder or Turn Away and Climb in the vast majority of the runs; these maneuvers were not always adequate and, in some runs, other maneuvers would have resulted in safer miss distances.
- The largest determinant of overall performance was the intruder's trajectory. Although subjects were effective at dealing with low convergence rate intrusions, the subjects' responses to high convergence rate intrusions generated the same, or worse, performance than no reaction at all.
- An alerting and avoidance system should be capable of recognizing the properties of an intrusion and reacting accordingly. Even with the display of convergence rate and trend information, subjects appeared to react to all intrusions in a very similar manner, resulting in many late alerts, missed detections, and incorrect avoidance maneuvers.
- Therefore, the design of an alerting and avoidance system should consider several factors. First, the late alerts of the subjects and the high workload involved with alert generation suggest the alerts should be generated by an on-board system. Second, the incorrect avoidance maneuvers selected by the subjects suggest the task of selecting an avoidance maneuver may also be a likely candidate for automation, or that alerts must be given with sufficient warning that the pilot preferred maneuvers, Turn Away or Turn Away and Climb, are safe. Finally, care must be taken that the pilot trusts the automatic alerts enough to follow automatically generated alerts; this experiment appears to show that pilots do not use the same alerting schema as a more efficient system would.

Appendix A. Debriefing Form

Debriefing

Age? _____ Gender? _____ Student? _____

Video Game Junkie? _____ Pilot? _____

Are You Involved in Research in The Following Areas:

Displays? _____ A/C Dynamics? _____ Collision Avoidance _____

With each display, how did you decide when to 'call an alert'?

Display 1: _____

Display 2: _____

Display 3: _____

Display 4: _____

How Do You Think the Turbulence-Induced Workload Affected Your

Ability to Monitor the 'Parallel Traffic'? _____

Comments on Sim?: _____

References

Ebrahimi, Y.S. "Parallel Runway Requirement Analysis Study" NASA Contractor Report 191549 Volume 1, December 1993

Owen, M.R. "The Memphis Precision Runway Monitor Program Instrument Landing System Final Approach Study" DOT/FAA/NR-92/11, May 1993

Pritchett, A.R., B.D. Carpenter, K. Asari, J. K. Kuchar & R. J. Hansman "Issues in Airborne Systems for Close Parallel Approaches" Proceedings of the 14th Annual Digital Avionics Systems Conference, Cambridge MA, October 1995

Pritchett, A.R. & R. J. Hansman "Preliminary Flight Simulator Evaluation of Human Factors Issues for Collision Avoidance During Close Parallel Approaches" ASL Report, In Progress

Shank, E.M. & Hollister, K.M. "Precision Runway Monitor" Lincoln Laboratory Journal, Volume 7, Number 2, 1994