AD-778 951

SIMULATOR EVALUATION OF PILOT ASSURANCE DERIVED FROM AN AIRBORNE TRAFFIC SITUATION DISPLAY. PHASE II. TRAFFIC AWARENESS IMPROVEMENT

David Melanson

Massachusetts Institute of Technology

Prepared for:

Federal Aviation Administration

July 1973

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.
FAA-EM-74-10

2. Government Accession No.

3. Recipient's Catalog No.
AD 778 951

4. Title and Subtitle
Simulator Evaluation of Pilot Assurance Derived from An Airborne Traffic Situation Display, Phase II - Traffic Awareness Improvement

5. Report Date
[illegible]

6. Performing Organization Code

7. Author(s)
David Melanson

8. Performing Organization Report No.

9. Performing Organization Name and Address
Massachusetts Institute of Technology, Man-Vehicle Laboratory, Electronic Systems Lab., Flight Transportation Lab., 77 Massachusetts Avenue, Cambridge Mass. 02139

10. Work Unit No.
011-001

11. Contract or Grant No.
DOT-FA-71 WAI 234

12. Sponsoring Agency Name and Address
Department of Transportation
Federal Aviation Administration
Office of Systems Engineering Management
Washington, D.C. (20591)

13. Type of Report and Period Covered
Phase II - Final Project Report; March 1, 197[illegible] - June 30, 197[illegible]

14. Sponsoring Agency Code
FAA-OSEM

15. Supplementary Notes
The work reported in this document was performed at the Massachusetts Institute of Technology by the three Laboratories listed above under Air Force Contract F19628-70-C-0230 with Lincoln Laboratory as per Interagency Agreement DOT-FA-71 WAI-234

16. Abstract

The research program described in this report evaluated the pilot assurance derived from an Airborne Traffic Situation Display (ATSD), assurance being defined in terms of the pilot's awareness of his traffic environment. Forty-four professional pilots participated in a number of simulation experiments designed to measure pilot awareness using today's party line voice communications and a possible future system employing the ATSD. Tests were run with and without an in-trail spacing task and with one and two-man flight crews. In addition, a pilot's ability to detect and react to conflict situations was measured during both single and parallel runway operations. The effects of conflict alarms and the frequency of updating information were also examined. All experiments were conducted in a fixed-based simulator configured as a Boeing 707.

The ATSD was found to be superior to the party-line communication channel as a source of traffic awareness.

With no spacing task, the detection of conflicts prior to the point of closest approach occurred in all cases employing the ATSD regardless of whether an alarm was used or whether a crew or single pilot was being tested. With a spacing task, a high percentage of the conflicts were detected by single pilots, but not always in time to take safe evasive action, particularly during closely-spaced parallel approaches. Tau (range divided by range rate) alarms reduce the reaction time of both crews and single pilots in responding to conflicts in some, but not all, conflicts.

17. Key Words
ATC Displays
Air Traffic Control (ATC)
Situation Displays
Collision Avoidance
Human Factors
ATC Distributed Management

18. Distribution Statement
Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Va. 22151, for sale to the public.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	241	$6.00

Form DOT F 1700.7 (8-69)

ACKNOWLEDGMENTS

The prior work of Thomas Imrich Robert Anderson, and Jack Howell on the ATSD underlies much of the research effort described in this report. In particular, some of the experimental data collected by Howell in Phase 1 of the FAA program has also been used in the statistical analysis presented here.

The conscientious and helpful guidance of Mr. Edgar Post, the project's technical monitor at FAA, OSEM, and his associate Mr. Neal Blake is hereby acknowledged with gratitude.

The project engineer at M.I.T. is Mr. Mark Connelly of the Electronic Systems Laboratory. Work at the Man-Vehicle Laboratory was under the direction of Professor Renwick Curry and at the Flight Transportation Laboratory, supervision was by Professor Robert Simpson. The cooperation and advice of the many professional pilots who have voluntarily participated in this program is acknowledged with sincere thanks. A major portion of this report was submitted to the M.I.T. Department of Aeronautics and Astronautics in partial fulfillment of the requirements for a Master of Science degree.

Preceding page blank

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

1.1 Executive Summary

The Airborne Traffic Situation Display (ATSD), as the name implies, is a concept for providing Advanced Radar Traffic Control System (ARTS) information in the cockpit, processed to place the equipped aircraft at the center of a heading-up display. Other aircraft are displayed with identification/altitude/ground speed data tags and in proper relationship to navigation routes, fixes, and the equipped aircraft. This executive summary lists the objectives of the U.S. Air Force one-year effort and the FAA two-year Phase I and II efforts, summary conclusions reached in analyzing both Phase I and II data, suggested additional simulation studies, and closes with suggestions to the reader who desires more details on various aspects of the ATSD evaluation program.

The objectives of the U.S. Air Force effort were to develop an ATSD-equipped flight simulator corresponding to the Boeing 707 jet transport, and use it to investigate the effects of an ATSD on safety, efficiency and capacity in the third generation ARTS Air Traffic Control (ATC) system environment and to optimize the display configuration and operational procedures by conducting basic tracking, ATC procedural, and spacing tests. A digital computer was used to solve the equations of motion of the subject aircraft, simulate movement of other aircraft in Boston's Logan International Airport terminal area, operate the ATSD and two sets of flight instrument displays. and process the experimental data on-line.

The objectives of the FAA Phase I continuing evaluation of the ATSD were to complete the optimization of the display configuration, to evaluate its use in a discrete address communication environment and in enhancing safety by enabling pilots to monitor traffic and participate more actively in normal and abnormal terminal area ATC operations, including 2,500 foot closely spaced independent, parallel runway operations. In addition, the evaluation was to be structured so as to determine whether the anticipated increase in pilot assurance, i.e., awareness resulting from having the ATSD in the cockpit was measurable, and what effect its presence would have on pilots being willing to accept 2,500 foot spaced parallel runway operations.

The objectives of the FAA Phase II ATSD evaluation were to determine if the "pilot assurance" value of the ATSD and the pilot's ability to detect gross system errors could be improved over the findings in Phase I by:

- Deleting the pilot's unfamiliar "inner loop" spacing task.

- Making the simulation more realistic by adding a second crew member.
- Providing a finer scale on the ATSD for the closely spaced phase of approaches.
- Increasing pilot simulator and ATSD familiarization and adding conflict detection training.
- Providing computer-generated alarms of potential conflicts.

In addition, the effect on pilot assurance and his ability to detect gross errors was to be determined when "radar and imperfect navigation noise" is added to other aircraft appearing on the ATSD.

The more important of the conclusions resulting from the analysis of Phase I and II simulation results and prior research appear below:

1. The ATSD is a positive aid to the pilot in establishing and maintaining separation during his approach to the outer marker, and the precision is better than with current methods.[5]

2. Delegating the final merging and spacing task to a pilot having the ATSD reduces controller-pilot communications and controller workload.

3. While not totally effective in assisting the pilot to detect and avoid blunders in terminal area operations, the ATSD is more than an adequate replacement for the voice party-line controller-pilot communication link.

4. The ATSD is a valuable aid to the pilot in flying in the terminal area and results in increased assurance because he is continually oriented during extensive radar vectoring and because he is aware of other aircraft operations in his vicinity.

5. Many pilots felt the ATSD gave them the confidence to fly independent 2,500 foot spaced parallel approaches.

6. Nearly all pilots noted an increase in workload due to the ATSD, but a large majority of those questioned on the subject felt that the benefits out-weighed the increased workload.

From the above, one can draw the general conclusion that although the ATSD is not envisioned as a mandatory device, it does offer advantages to pilots flying in high-density areas and, therefore, the simulation program should be continued to:

- Improve its utility in detecting and avoiding worst case blunders in independent parallel runway operations.
- Explore its application to airport surface navigation and control.
- Determine the reduced longitudinal separations pilots would find acceptable under Instrument Meteorological Conditions (IMC) if the preceding aircraft can be seen clearing the runway on the ATSD and assuming wake turbulence is not the limiting factor.

- Explore combining digitized weather radar information on the display to improve the pilot's ability to navigate around severe weather.

Readers desiring more detailed information will find additional background in section 1.2 and a more extensive discussion of the conclusions in section 1.5 and in Chapter 5. More details of the Phase I effort will be found in section 1.3. Phase II is discussed in sections 1.4 and 5.1, and in Chapter 2. The ATSD format and the cockpit simulation facility are treated in sections 1.6 and 1.7 respectively. A postulated complete ATSD system is described in section 1.8. Chapters 3 and 4 and the Appendix go into great detail on the statistical analysis of the data and will be useful primarily, to human factor methodologists. Chapter 5, however, will be of interest to those looking for more information on the Phase I and II results and on the conclusions that are presented in this first section. Unless the reader has read earlier reports on the ATSD, it is recommended that he read Chapter 2 before studying Chapter 5.

1.2 Background

The basic idea of presenting an onboard pictorial display of traffic information is not new. This was introduced as early as 1946 by RCA in their TELERAN program.[1] In 1963, the FAA conducted simulations using a cockpit display (Sluka, 1963),[2] and more recently in 1966, an effort to provide televised radar pictures for pilots was tested in the Boston area under the direction of the FAA.[3] These prior investigations showed that definite advantages for the air traffic control (ATC) system could be derived from such information, but that the attention span required to derive enough information from a rather poor quality of display was too high. What these initial efforts lacked was the ability to provide the essential elements of information about traffic in an easily discernable format for quick interpretation by the pilot. The advent of computerized radar tracking systems in the terminal areas and computer generated displays now overcomes this previous drawback with symbolic and alphanumeric presentation of traffic information in an appropriate format.

One approach to the presentation of this information is presented by Bush et al (1970)[4] where it is assumed that an Airborne Traffic Situation Display (ATSD) could be devised that would make portions of the NAS/ARTS (National Airspace System/Automated Radar Terminal System) computerized data base available to the air crew by an omnidirectional broadcast of traffic information throughout the terminal area. This information would be received by all aircraft equipped with an ATSD and onboard processing would be performed to present a symbolic representation with either a north-up or heading-up display format.

In July of 1970, under U.S. Air Force contract, the M.I.T. Lincoln Laboratory initiated a one-year effort by a consortium of on-campus research groups consisting of the Electronic Systems Laboratory, the Flight Transporation Laboratory, and the Man-Vehicle Laboratory to evaluate the potential usefulness of displaying selected ARTS data in the cockpit and the effect that the availability of such information would have on terminal area procedures and capacities. The ATC functions that could beneficially be delegated to the cockpit were to be specified, as well as the optimum way of utilizing airborne displays in implementing these functions. To carry out this program, the consortium developed an Air Traffic Situation Display (ATSD)-equipped fixed base flight simulator corresponding to the Boeing 707 jet transport. An Adage AGT-30 digital computer was interfaced with the cockpit and used to solve the equations of motion of the subject aircraft, simulate the movement of other aircraft in the terminal area, create the ATSD and two flight instrument displays, and process experimental data on-line.

Tests were run on the simulation facility to evaluate the operational value of the ATSD in a realistic environment modeled on the airspace structure of Boston's Logan International. Typical tasks were sequencing, merging, spacing in trail, following route structures, and conveyor belt tracking. A basic evaluation of the concept was completed by Imrich[5] in 1971 with initital indications that the ATSD could increase spacing accuracy and decrease response time in emergency conditions. In addition, the ATSD extended the senses of the pilot to permit nearly Visual Meteorological Condition (VMC) operational approach rates under Instrument Meteorological Conditions (IMC) and reduced radar controller workloads significantly by increasing pilot participation in ATC functions. Other studies by Anderson[6] examined the effect of different display formats on pilots' scan workload and ability to follow other aircraft in trail.

Following the one-year program for Lincoln Laboratory, the Office of Systems Engineering Management of the Federal Aviation Administration sponsored a six-month Phase I investigation of the use of the ATSD as a traffic monitor in busy terminal areas to increase pilot assurance. Twenty professional pilots were exposed to a set of typical normal and abnormal terminal approach situations. Three basic communications-display modes were employed: party-line and no ATSD, party-line with ATSD, and discrete address with ATSD. The level of pilot assurance was determined by their detailed knowledge of each situation as measured in stop-action quizzes and by their ability to detect conflicts.

Workload or the degree of difficulty the pilots experienced in acquiring relevant information about the situation was also regarded as a component of assurance. Specific problem areas emphasized in the test scenarios were simultaneous approaches to closely-spaced parallel runways, blunder detection and resolution, and providing a "picture" for the pilot when discrete address data links replace current ATC party-line communications. All of the Phase I tests employed single subjects and involved an in-trail spacing task. No conflict detection training was given to this group of subjects prior to testing. Report FAA/EM-72-3 by Howell documents the results of the Phase I effort.

An extension of this work, designated as Phase II, has investigated the effects of adding a second crew member, providing computer-generated alarms of potential conflicts, pilot training in conflict detection, and using one second instead of four second position updates. In addition, the in-trail spacing task was eliminated in the Phase II test series and irregularities added to target motion to make the scenarios more realistic. Only the discrete address communications mode with an ATSD were employed in these tests. A side-task measure of crew perceptual workload was also included in supporting experiments. Twenty-four professional pilots participated in the Phase II tests.

1.3 Summary of Phase I Work

This section is based on and supplements the background given in section 1.2.

Whether or not the pilots can maintain separations with an ATSD depends largely on the type and quality of information that the pilots have about their position relative to navigation route segments and fixes and other aircraft. To evaluate the information transfer, Howell (1972)[7] in Phase I undertook a set of simulation studies to determine the type of information that the pilot has with respect to his navigation and traffic situation within the terminal area. The information the pilot has about his own aircraft and those surrounding him would seem to be of significant interest, and to our knowledge, this was the first time that a quantitative study has been undertaken of the subject. Phase I evaluated the information transfer with four combinations of the following display/communication factors:

ATSD vs. no ATSD

Party-Line vs. Discrete Address Communications

Each factor has two levels. A combination of one level from each factor is called a treatment and all four factorial combinations were considered. The combination of no ATSD and party-line communications corresponded to the present day air traffic control system and provided a convenient baseline for comparison purposes. No display, discrete address commands corresponded to the case where the pilot received only the commands directed towards his aircraft and none of the other commands were heard. This will be the situation that is presently envisaged with the Discrete Address Beacon System (DABS). Traffic display, discrete address commands corresponded to the situation where a display is added to a discrete address communication system, whereas traffic display, party-line communications corresponded to the information situation that would exist if the traffic display were added to the present day air traffic control system.

The original request of the sponsoring agency was the evaluation of the ATSD as an "assurance" device. To do this, assurance was equated with awareness. Although it is true that the majority of pilots would like to have more information about the traffic situation as it evolves around them, there probably is a sub-population of pilots who would be more assured by not having to assimilate this type of information and would like to rely on the air traffic control system completely. Awareness was defined to consist of the following elements:

1. The pilot's knowledge of his current position with respect to the air route structure.

2. The pilot's knowledge of the position of other aircraft around him.

3. The pilot's ability to predict the evolution of the traffic situation in the short term (especially the evolution of abnormal situations)

4. The pilot's ability to choose appropriate escape maneuvers should an emergency occur.

The Phase I simulations, then, were designed to evaluate these facets of awareness under the four display/ communications conditions described above.

An analysis of this data indicated an increased information transfer when the ATSD was employed, but also showed that a higher workload level on the part of the pilot was required. In addition, a significant, although not entirely satisfactory, increase in the ability to detect conflicts was observed.

The results from these tests were confounded by the fact that, when the ATSD was used, the pilot was also required to perform a spacing task, that is, to follow the preceeding aircraft in trail by a specified distance. The inclusion of this task was thought to distort the information transfer for the ATSD modes with both discrete and party-line communications by increasing the pilot's workload. More specifically, the spacing task was thought to focus the subject's attention on the aircraft that was being followed with a corresponding decrease in attention to other aircraft.

All tests conducted in Phase I used a single pilot performing the functions that would normally (at least in the air carrier case) be performed by a two or three man crew. For this reason the pilot's ability to monitor the ATSD for traffic awareness and conflict detection functions was believed to be less effective than would be possible with a more realistic crew situation because the overall workload imposed on the pilot was considerably increased.

1.4 Scope of the Phase II Work

This work is a direct extension of the work done in Phase I. Near-terminal area simulations were conducted to measure pilot awareness using both single pilot and two man crew simulations. Greater emphasis in Phase II was placed on conflict detection measurements particularly during independent operations on closely spaced parallel runways. In light of the data obtained in Phase II, a further analysis of portions of the Phase I results was performed.

This report addresses itself to a comparison of pilot awareness with variations in five major factors:

(1) alarm vs. no alarm

(2) one second update vs. four second update

(3) spacing task vs. no spacing task

(4) one-man crew vs. two-man crew

(5) today's ATC system (no ATSD, party-line communications) vs. future ATC system (ATSD, discrete address communications)

Each factor has two levels. A combination of one level from each factor is called a treatment. It is not a factorial experiment in that all possible combinations of factor levels were not considered. The differences in information transfer and conflict detection are examined only for those treatments of practical interest.

The spacing task comparison consists of either including or excluding the in-trail spacing task. Only discrete communication with an ATSD was utilized in these tests.

The two flight crew options considered employed either a single pilot or a two man flight crew. These tests were conducted with a discrete communication channel and an ATSD. The subject pilots were not required to perform an in-trail spacing task.

The effects of position update rate (either four second or one second) on the conflict detection capabilities of single pilots under the discrete communication with ATSD display/communication format was tested during simulations of independent operations on closely spaced parallel runways. However, the effect of a proximity alert and emergency alarm on conflict detection was examined during both single runway and parallel runway operations.

Pilot opinions concerning awareness, workload, and simulation realism were solicited at the conclusion of the experiments via a questionnaire. The responses to specific questions, as well as pilot comments, are presented in this report.

1.5 Phase I and II Conclusion Highlights

Conclusions are stated somewhat differently in this section than in the Executive Summary in order to expand on the Executive conclusions by associating supporting conclusions with primary conclusions. These Phase I and II conclusion highlights provide a condensed version of the Chapter 5 results and conclusions. Primary conclusions are preceded by numerals and supporting conclusions by small case letters:

1. With the ATSD, a pilot can consistently space his aircraft more accurately behind the preceding aircraft at the outer marker than a controller can who is sequencing and spacing a number of aircraft. In addition, delegating the spacing task to the pilot reduces pilot-controller communications.[5,7]

   a. The presence or absence of a spacing task does not affect the pilot's estimation of information components in a statistically significant* way, but the spacing task does increase the percentage of null (no answer) responses in the stop-action quizzes.

b. In stop-action quizzes, subject pilots made fewer "gross errors" with the ATSD than they do with party-line information. A "gross error" in this report is defined as assigning the wrong sequence to a target or having it originate at the wrong feeder fix.

2. The ATSD, unaided by intruder alarms, permits a high percentage of conflicts to be detected, but not always in time to take evasive action, particularly during closely spaced parallel approach operations.

---

*In evaluating results, it is important to understand the term "statistically significant." A statistically significant result is one that would be highly unlikely to occur by chance with the data as given. In this report, one chance in twenty (5%) has been arbitrarily selected as the threshold level of statistical significance. With a limited number of subjects in each test population, it is common to measure substantial differences in the results from two treatments, but still not be able to classify the differences as statistically significant because of the small number of samples. Conversely, with a large number of samples, a small difference in results may be statistically significant, but be of no practical consequence. To assist those readers who have no background in statistical analysis, a short explanation of the techniques used in this report is presented as an Appendix.

a. Crossover Tau (range divided by range rate) alarms reduce conflict detection times in some, but not all, conflicts.

b. The 28 second Tau alarm used in the closely-spaced parallel runway cases did provide a marginal degree of safety for the 18° banked turn crossover intrusion used in the simulations. This Tau was too high, however, for practical use in the real world because it probably would result in too many false alarms.

c. The reaction time of crews to a conflict situation tended to be somewhat longer than that of single pilots, both with and without alarms. The alarm reduced reaction times in Situation 6 (ILS acquisition blunder), but not in Situation 7 (ILS intruder crossover).

d. With no spacing task, the detection of conflicts prior to the point of closest approach occurred in 100% of the cases employing the ATSD in Phase II. These cases used both single pilots and crews, alarms and no alarms, and single and parallel runway situations. Howell's Phase I data, with the spacing task included, showed six missed detections in 32 conflict cases, hence the spacing task, at least for single pilots without alarms, seems to detract from the conflict detection performance. This result may be due in part to the more intensive training in conflict detection given to Melanson's Phase II subjects, but there is no way to isolate the effects of training and the spacing task with the present data base. Similarly, with 100% detection in all the Phase II cases, it is not possible to draw statistically significant conclusions with respect to the value of the alarm, the second crew member, or the target position update rate.

3. The Airborne Traffic Situation Display (ATSD) with discrete addressed voice communications is superior to today's party-line voice communications as a source of information about other traffic. Hence, it would be a more than adequate replacement for the voice party-line as a source of pilot assurance and awareness.

a. The accuracy of pilot estimates of information components in the stop-action quiz (target position, spacing, altitude, heading, and ground speed) depends on the situation (scenario), on the sequence of a particular target relative to ownship, and on the information component being measured.

4. Many pilots were of the opinion that closely spaced parallel runway independent operations might be acceptable with the ATSD.

5. Pilot opinion of the ATSD was generally favorable. Awareness with the display was superior to that achieved with the voice party line. Their confidence in being able to detect and resolve blunders with it was high, but nearly all pilots noted an increase in workload due to the ATSD. This subjective workload opinion was verified by the results of the perceptual side task tests.

6. Crews generally estimated the information components more accurately than single pilots, but the margin was not great enough to be classified as statistically significant. In addition, the crews had a smaller percentage of null (no answer) responses in the stop-action quiz and rated the ATSD higher in the opinion questionnaire.

1.6 Display Format

The ATSD was presented in the cockpit on a cathode ray tube (CRT) masked to a 7 inch square size. The CRT was mounted above the throttle pedestal where the weather radar is normally located in a Boeing 707. The display presentation was a heading-up, own-ship-centered format with a four or one second display information update rate. The display orientation, therefore, corresponded to the pilot's view of the external world. Traffic elements (i.e., other aircraft) are shown as small circles with dots at the centers.

Each element was trailed by three tracer dots that marked the past positions of that aircraft 12, 24, and 36 seconds previously. Associated with each traffic element was the NAS/ARTS data block showing aircraft identification, altitude in hundreds of feet, and ground speed in knots. The own-ship data tag had only a ground speed readout. Also displayed on the CRT were navigation stations, route structure, and ground features. This information provided the pilot with a pictorial display of his geographic position. A picture of the ATSD format is shown in Figure 1.6.1.

The display controls were mounted to the left of the CRT. These controls allowed the pilot to select the desired traffic and map information by adjusting the volume of displayed airspace and by limiting the alpha-numeric readout items for each aircraft. The major controls were, first, control of the altitude layer above and below the subject aircraft within which traffic would be displayed and, second, a control of horizontal range from ownship to the top of the CRT frame (4, 8, 16, 32, 64 and 128 nautical mile ranges were available).

The presence of the alpha-numberic readout items was controlled by four toggle switches, which could selectively eliminate identification, altitude, ground speed, and the tracer dots associated with the other aircraft. These display

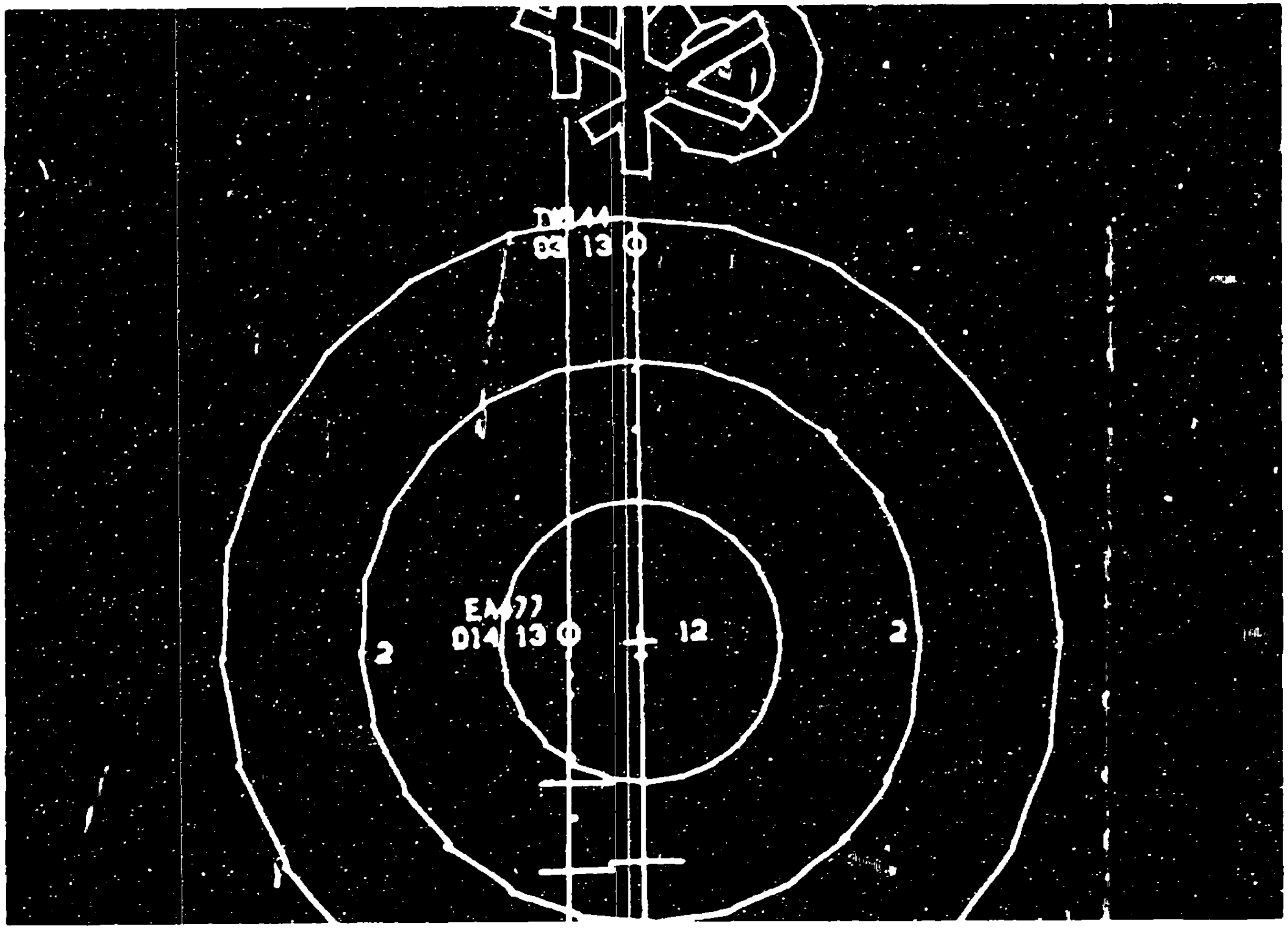


Figure 1.6.1 Airborne Traffic Situation Display Format

controls allowed the pilot to minimize clutter and to have high resolution in the areas of close aircraft spacing.

1.7 Simulation Facility

The basic component of the simulation facility is a fixed-base cockpit simulator that uses three cathode ray tubes to produce the ATSD and the primary flight instruments for both the pilot and first officer. The basic cockpit was built from an SST pre-prototype donated by the Boeing Company. The interior panels, switches, and instrumentation facsimiles are representative of a Boeing 707-123B aircraft. An ADAGE AGT-30 digital computer with a 16K core memory and a two microsecond cycle time was used to simulate the aircraft dynamics and perform calculations for the displays. An interior view of the cockpit simulator is shown in Figure 1.7.1.

The aircraft dynamics are representative of a Boeing 707 aircraft; the flight instrument package is patterned after the Collins FD-109 integrated flight system, but does not have the flight director functions. The flight control system simulates control wheel steering, an attitude rate command system that is available on the newer, wide-body jets. This not only provides a uniform flying workload while maintaining or changing altitudes, but it is felt that the attitude control task with control wheel steering

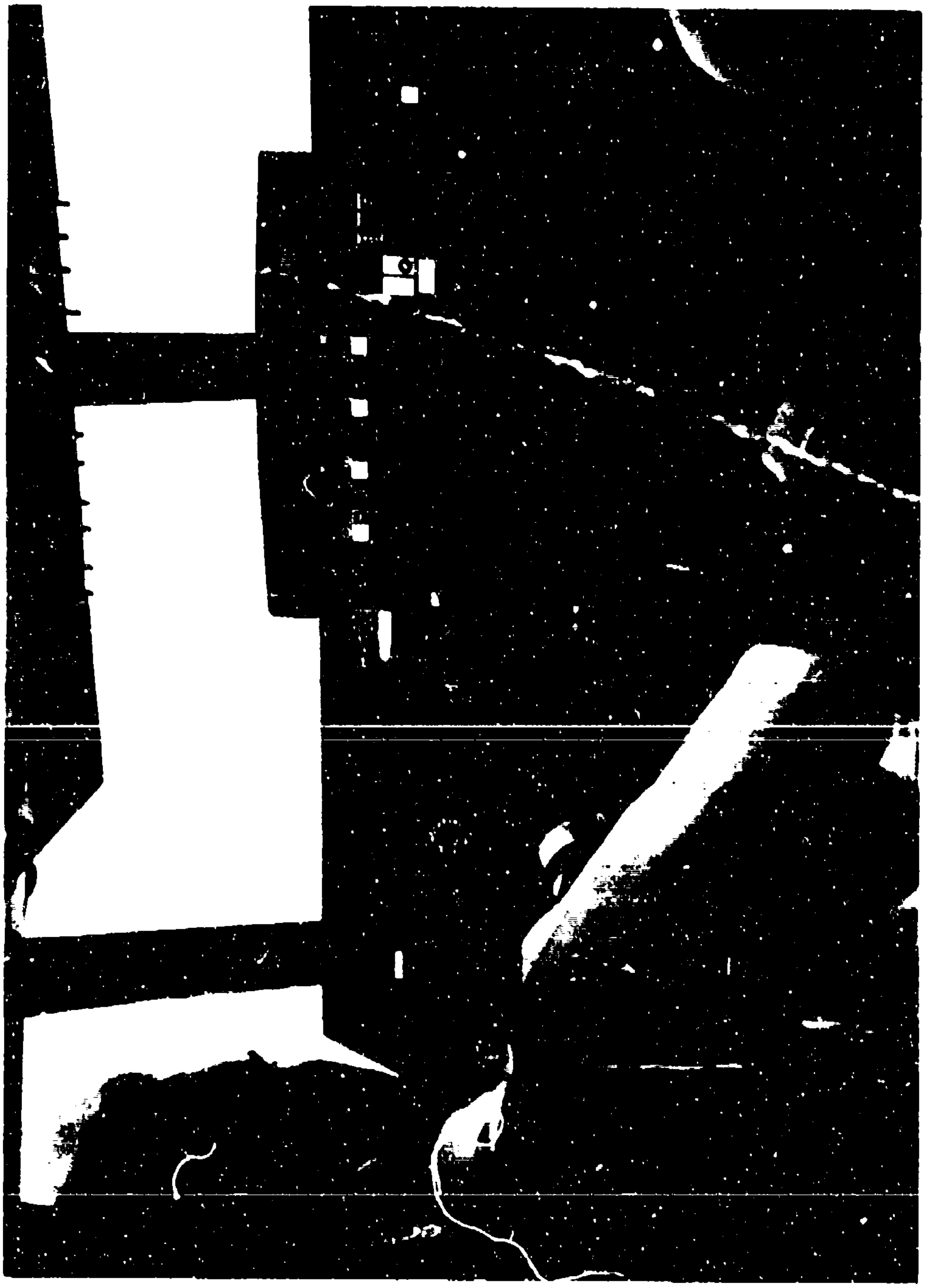


Figure 1.7.1 ATSD Simulator Cockpit Interior

in a fixed-base simulator (no motion cues) is comparable to that with conventional controls in a moving-base simulator (with motion cues).

The experiments were conducted with a simulated air traffic controller in an adjacent room. Communications between the subject-pilot and the controller were accomplished through the use of standard head sets (hot mike) and intercom lines. Live communications were used between these two stations. Responses from other aircraft in the traffic scenarios were stored sequentially on a tape recorder and played back, as required, in response to the controller commands.

The cockpit, controller display, computer, and associated interface hardware are shown in block diagram form in Figure 1.7.2.

1.8 Postulated System Configuration

A functional diagram showing one suggested system configuration proposed for a NAS/ARTS based cockpit Traffic Situation Display is shown in Figure 1.8.1.

Primary and beacon surveillance radars provide basic data to the computers in the air route traffic control centers and approach control centers. This basic data is processed along with flight plan information and used to generate the ATC controllers' displays. With limited reprocessing and

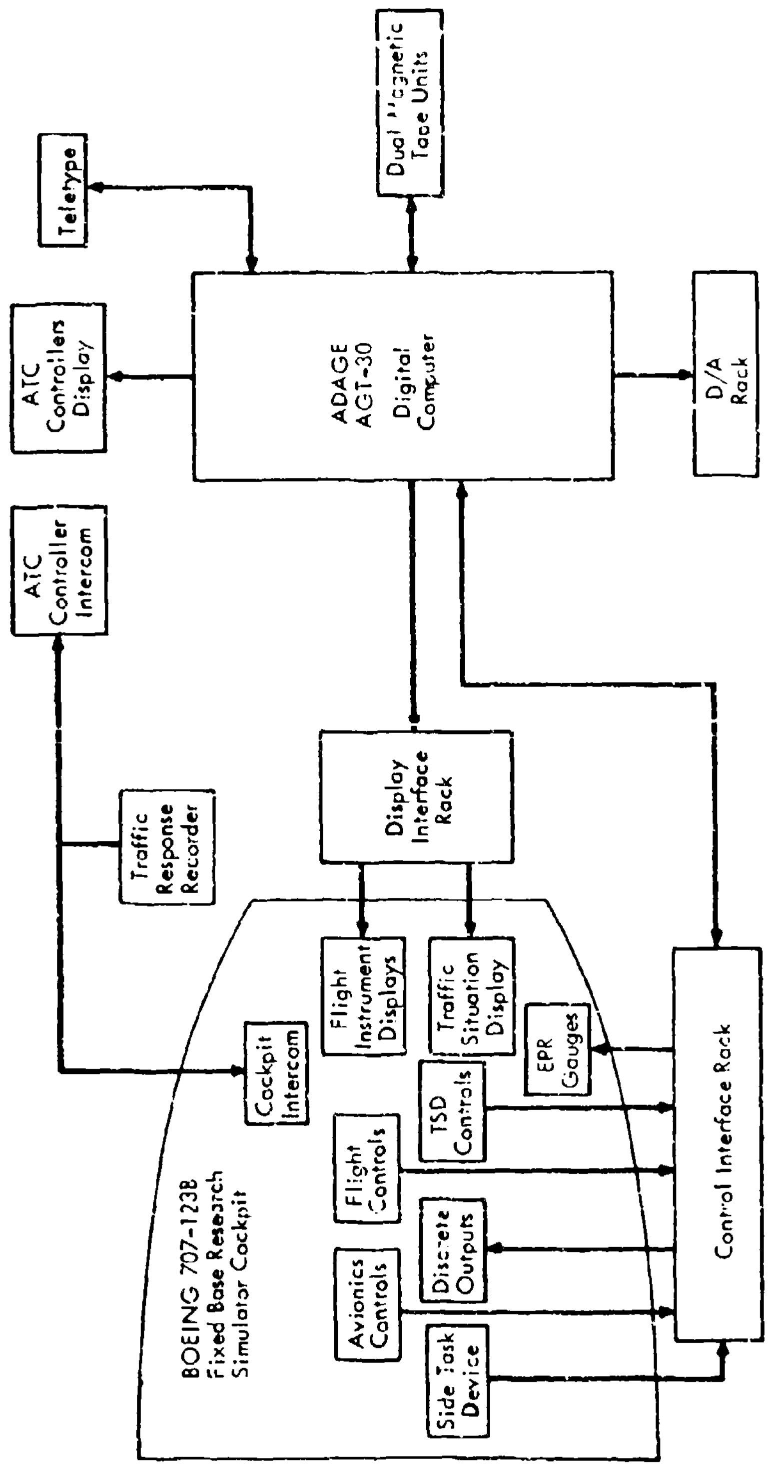


Figure 1.7.2 ATSD Simulation Facility Block Diagram

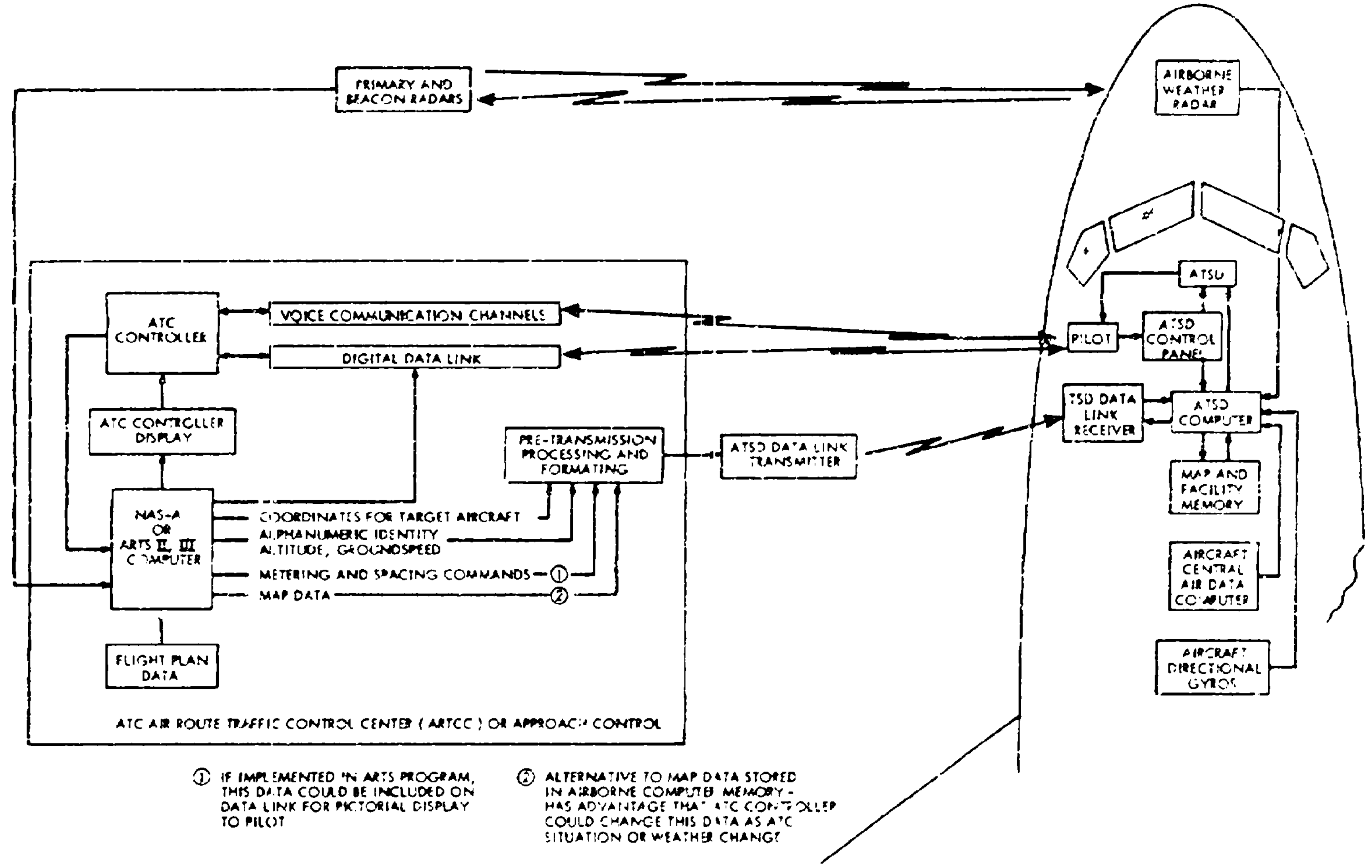


Figure 1.8.1 Airborne Traffic Situation Display System Diagram

formulating, portions of this basic data could be broadcast on a common radio frequency to aircraft within the facility control area to provide the data base for the ATSD. Transmission could be via a VHF digital data link. Studies indicate that the data required to service 100 aircraft in a terminal environment could be handled by an 8 kilobit/second transmission rate.[4] This would provide a complete traffic picture, including aircraft positions, identifications, altitudes (for mode C beacon equipped aircraft), selected map information, ground speed, and ground weather radar contours (if desired) every four seconds. This data rate can be accommodated in a 25 KHz VHF channel.

Aircraft equipped to receive the data link broadcast would use a small, on-board computer to select appropriate information from the data stream, process the data according to the settings on the pilot's display controls, and generate the ATSD picture. Other inputs to the ATSD computer would be:

1. Aircraft heading from the directional gyro would be used to orient the display to heading up.

2. Central air data computer (CADC) signals would provide own aircraft's altitude.

A study would be required to determine if background map data, which would not change frequently, could be economically carried in storage on board the aircraft.

The display of air-derived or ground-derived weather radar system information, processed and displayed on the ATSD at the appropriate range scales, also deserves further study.

CHAPTER 2

EXPERIMENTAL METHOD

The Phase II set of experiments were designed to compare the abilities of single pilots and flight crews in monitoring the ATSD for traffic information and conflict detection. In addition, the effect of alarms and information update rates on conflict detection was tested. Also a more detailed understanding of pilot awareness derived from an ATSD with no in-trail spacing task requirement was sought.

Part of the data from these tests was compared to portions of the data gathered in Phase I to eliminate some of the confounding of variables caused by the spacing task.

For this reason, there are really two experiments and, therefore, two experimental plans being discussed here. The experimental plan used in Phase I is recorded in Howell's report[7] and it was briefly discussed in section 1.3 of this report. The experimental plan used in the Phase II experiments is outlined in this chapter. An analytical plan used to analyze data covering aspects of both experiments is presented in the next chapter.

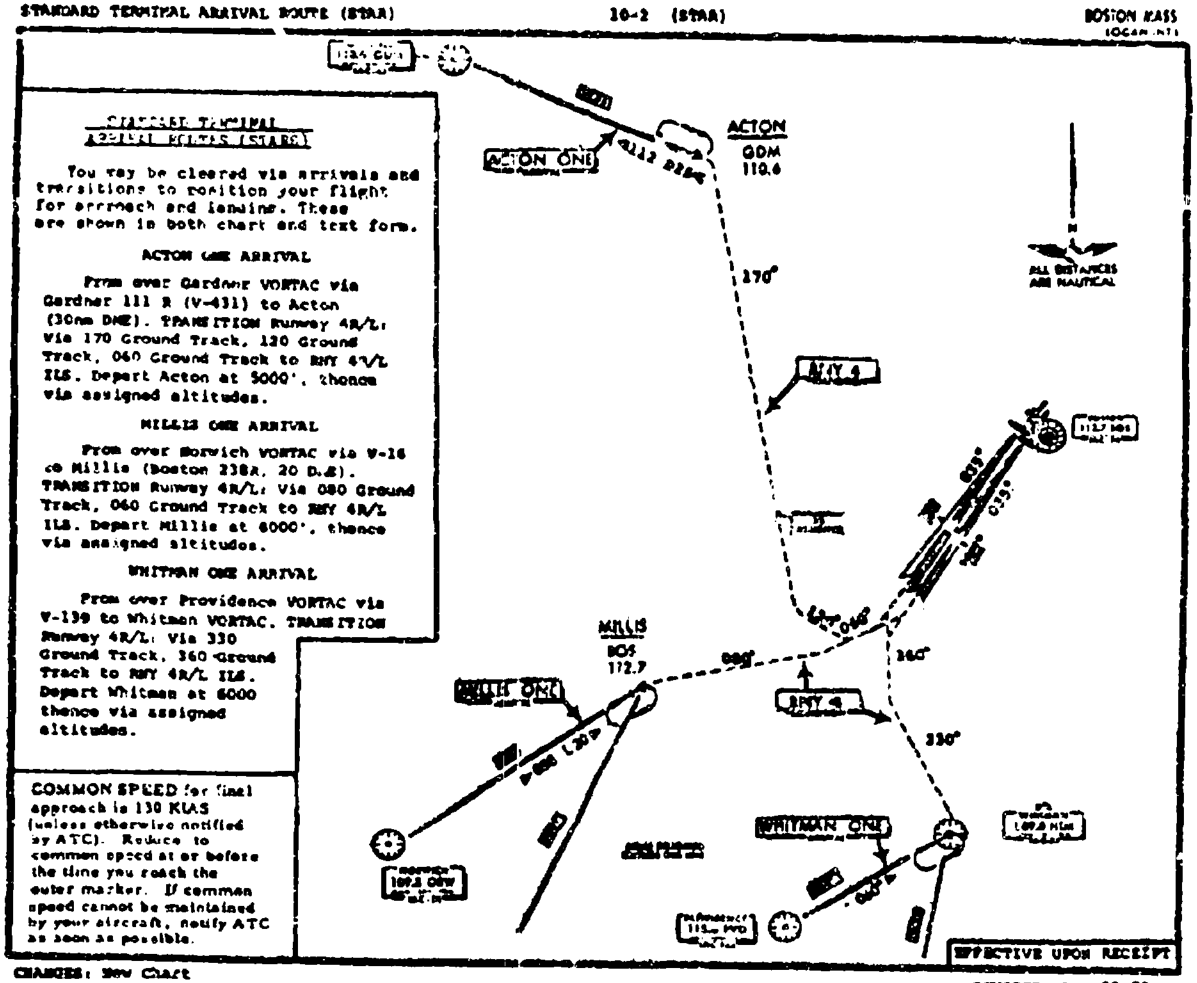
2.1 Scenario Development

In each of the seven basic traffic simulations, all the aircraft conformed to the Standard Terminal Arrival Route (STAR) chart constructed for this work. The transition routes appeared on the ATSD along with the three fixes and the two ILS (Instrument Landing Systems) courses for runways 04 left and 04 right (Figure 2.1.1). The situations were numbered sequentially from 1 thru 7.

Each data run began with a formal clearance which included a weather summary read by the air traffic controller. The simulation was begun once the correct response was given. The subjects were then guided by radar vectors except in the spacing task treatment where the subjects used the STAR structure as a nominal course. When the party-line communication channel was employed, the controller read a series of commands intermittently directing the pre-programmed targets. The commands were timed to fit the pre-programmed trajectories, and were sequenced by referring to a stopwatch. Responses from the program targets were played back from a tape-recorder, while the dialogue with the subject-pilot was, of course, live.

Four of the seven scenarios (numbers one thru four) consisted of merging streams of traffic to a single active runway, while the remaining three consisted of independent approaches to closely spaced parallel runways. The runway centerline separation was 2600 feet.* Except for one blunder situation (situation 6), a vertical separation of 1000 feet was maintained at the turn-on points for the two ILS's. In-trail separations of three to four miles were maintained

*Closely-spaced parallel runways are nominally taken 2500 feet apart. The extra 100 feet has no major effect.



Format courtesy: Jepco Avigation

Fig. 2.1.1 Simulated Boston STAR Chart

under the direction of the simulated approach controller.

2.2 Stop-Action Quiz

One of the primary goals of this investigation was to determine the type of information the simulator pilot has about other traffic within the terminal area. A stop-action quiz was used to evaluate this type of information. When the situation had developed to the extent that a reasonable amount of information had been presented to the pilot and the traffic density was approaching maximum, the simulation was halted without warning. A given situation was always halted at the same point for all subjects. Presentations on all CRTs were blanked. The pilot or crew was then required to complete the quiz on a map shown in Figure 2.2.1. Subjects were asked to supply the following information about each aircraft in the traffic situation: position, identification, landing sequence number, heading, altitude, and ground speed. Stop-action quizzes were completed for single runway simulations only. These maps were the primary source of quantitative information in the results reported here. Figures 2.2.2 thru 2.2.5 show the ATSD at the time of the stop-action quiz. Figure 2.2.6 shows a typical completed quiz map.

The stop-action quiz responses were graded for accuracy and completeness. Errors in subject estimates of information components were recorded. Those components which

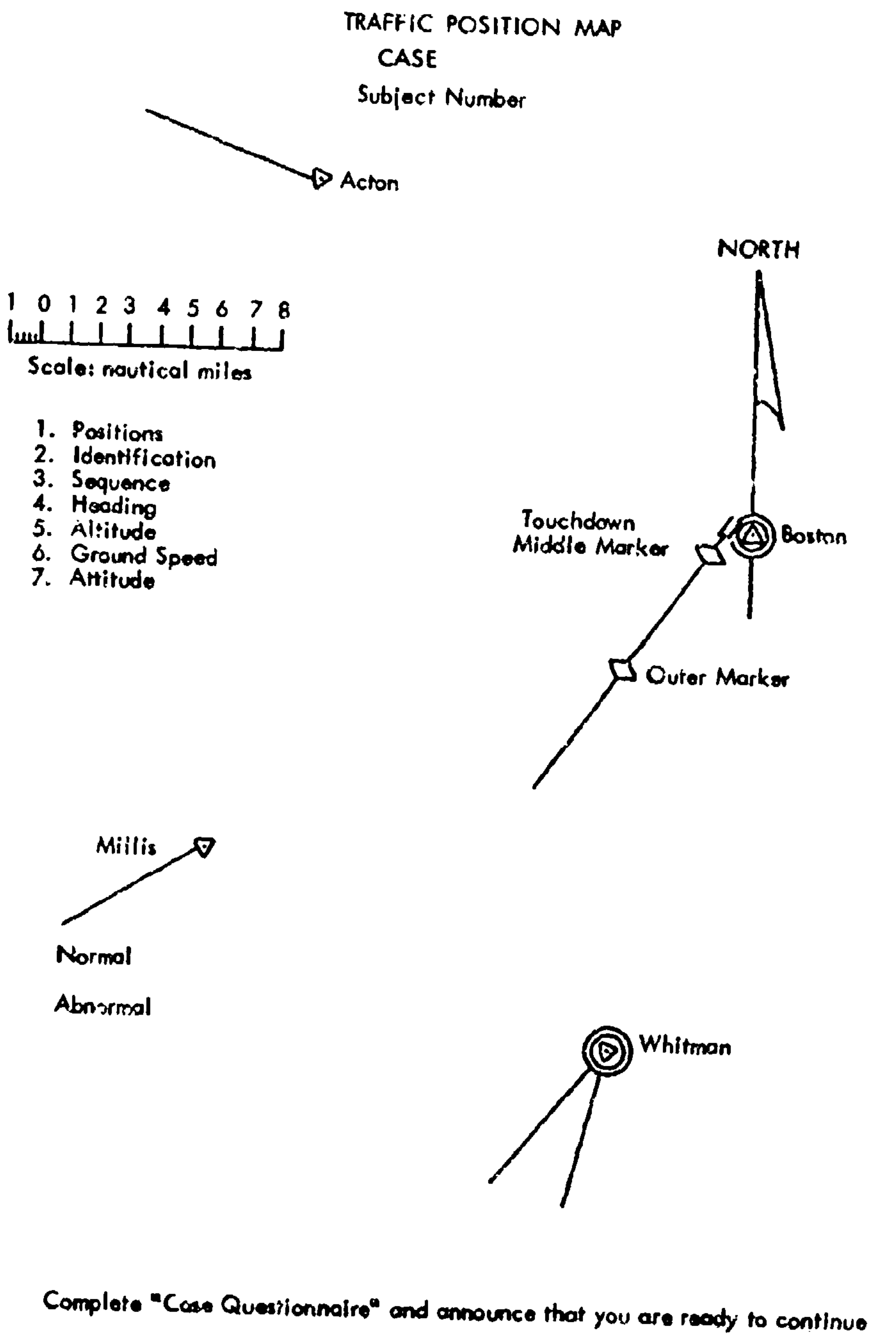


Fig. 2.2.1 Map used in Stop Action Quiz

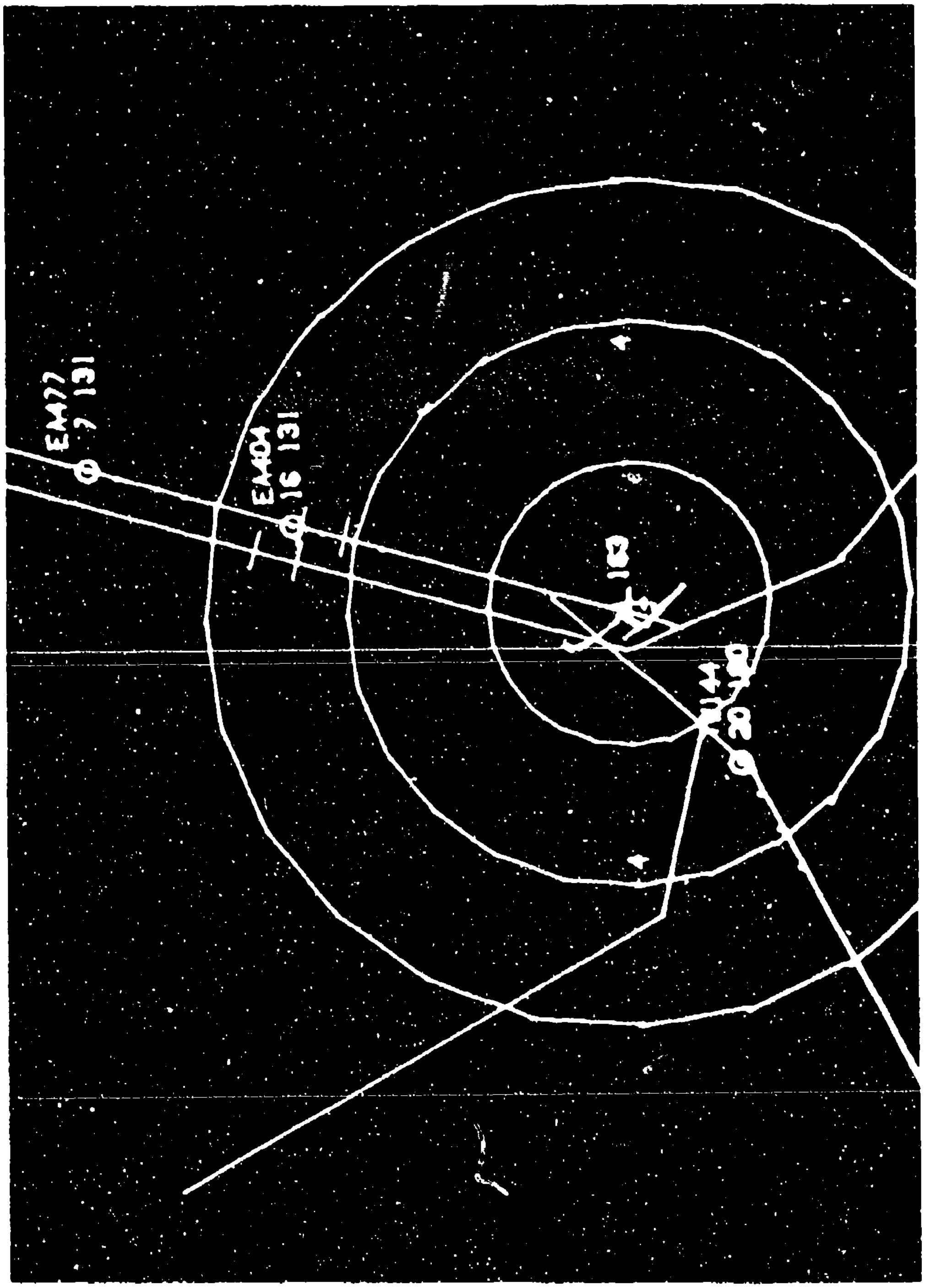


Figure 2.2.2 ATSD Presentation at Stop-action Quiz for Situation 1

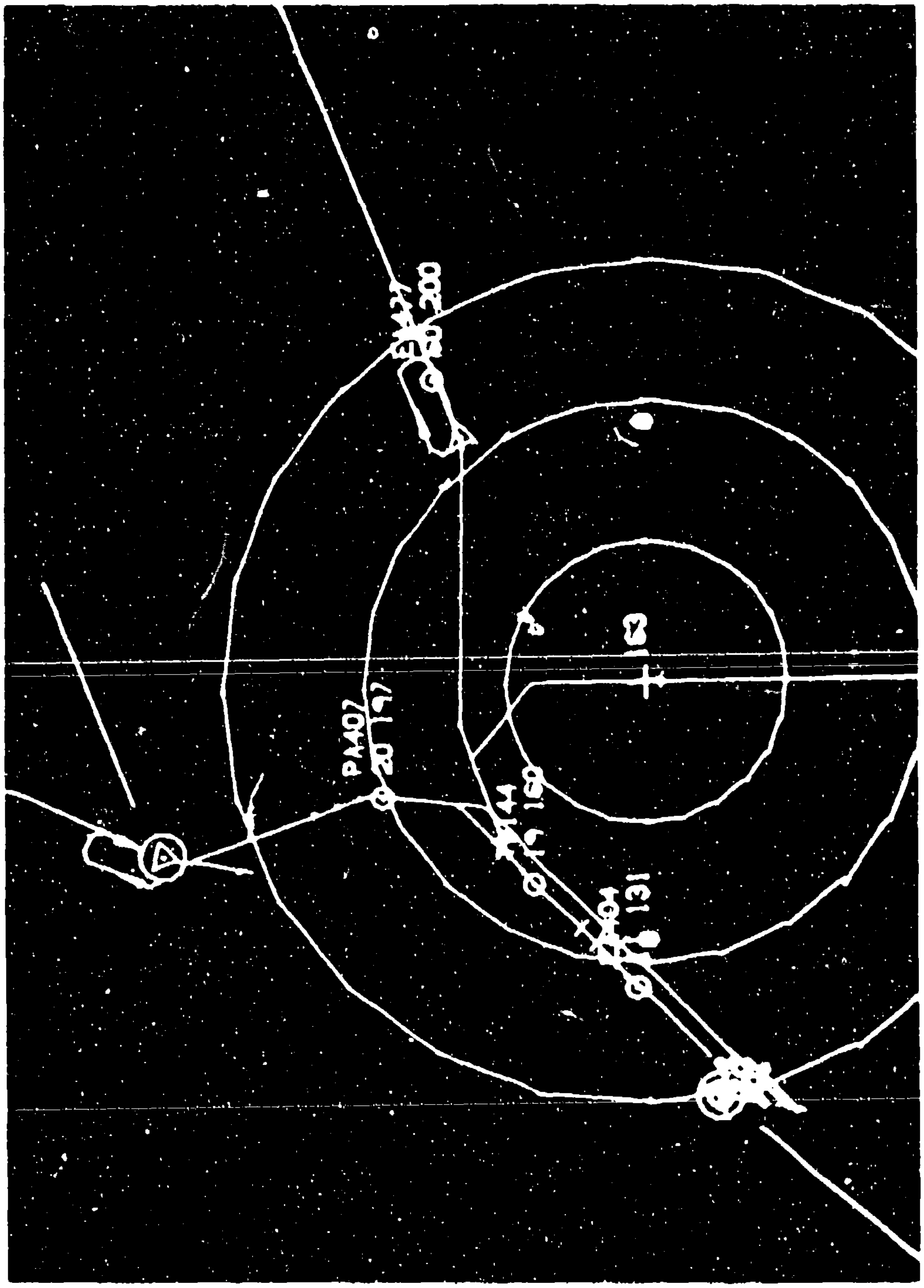


Figure 2.2.3 ATSD Presentation at Stop-action Quiz for Situation 2

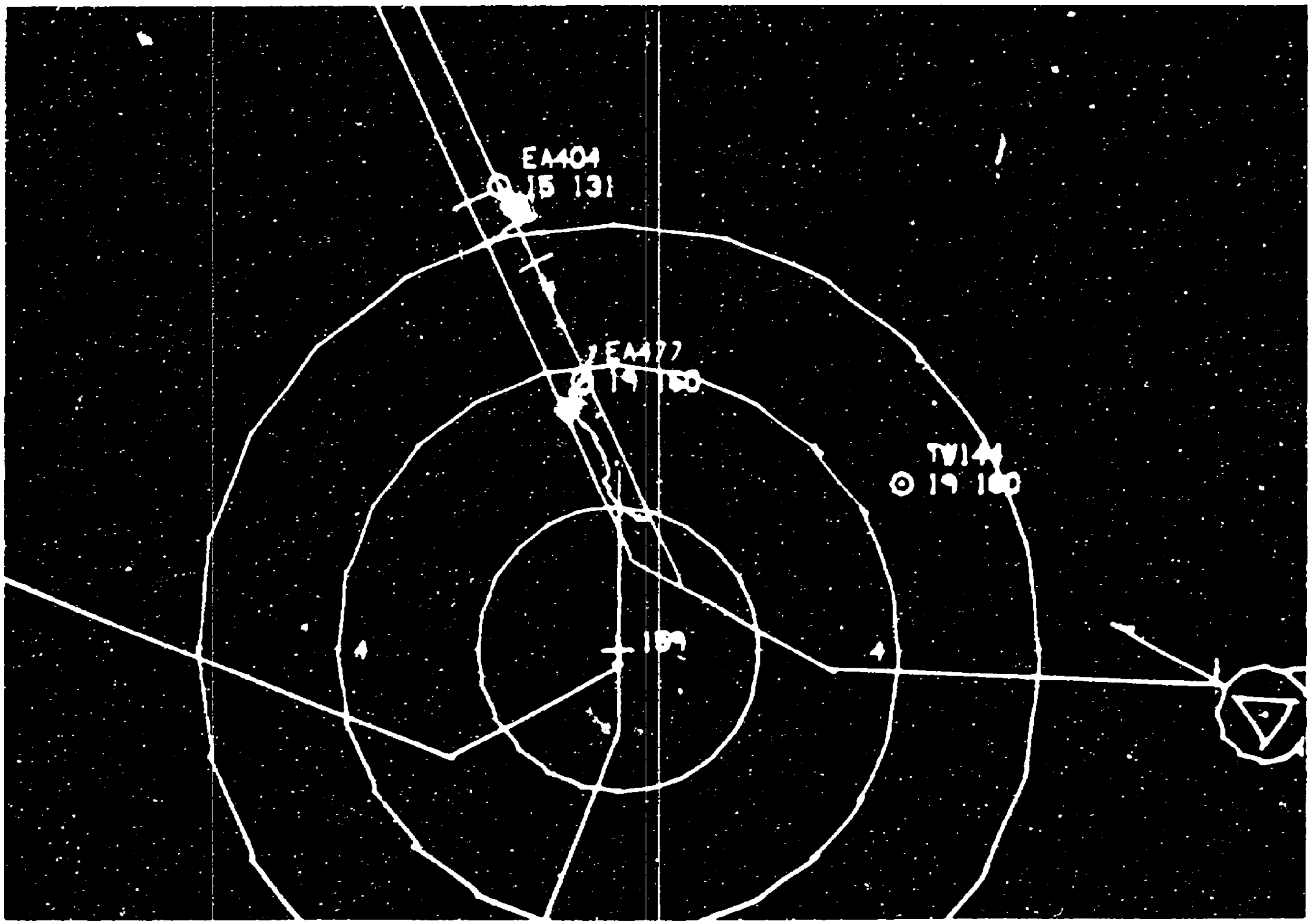


Figure 2.2.4 ATSD Presentation at Stop-action Quiz for Situation 3

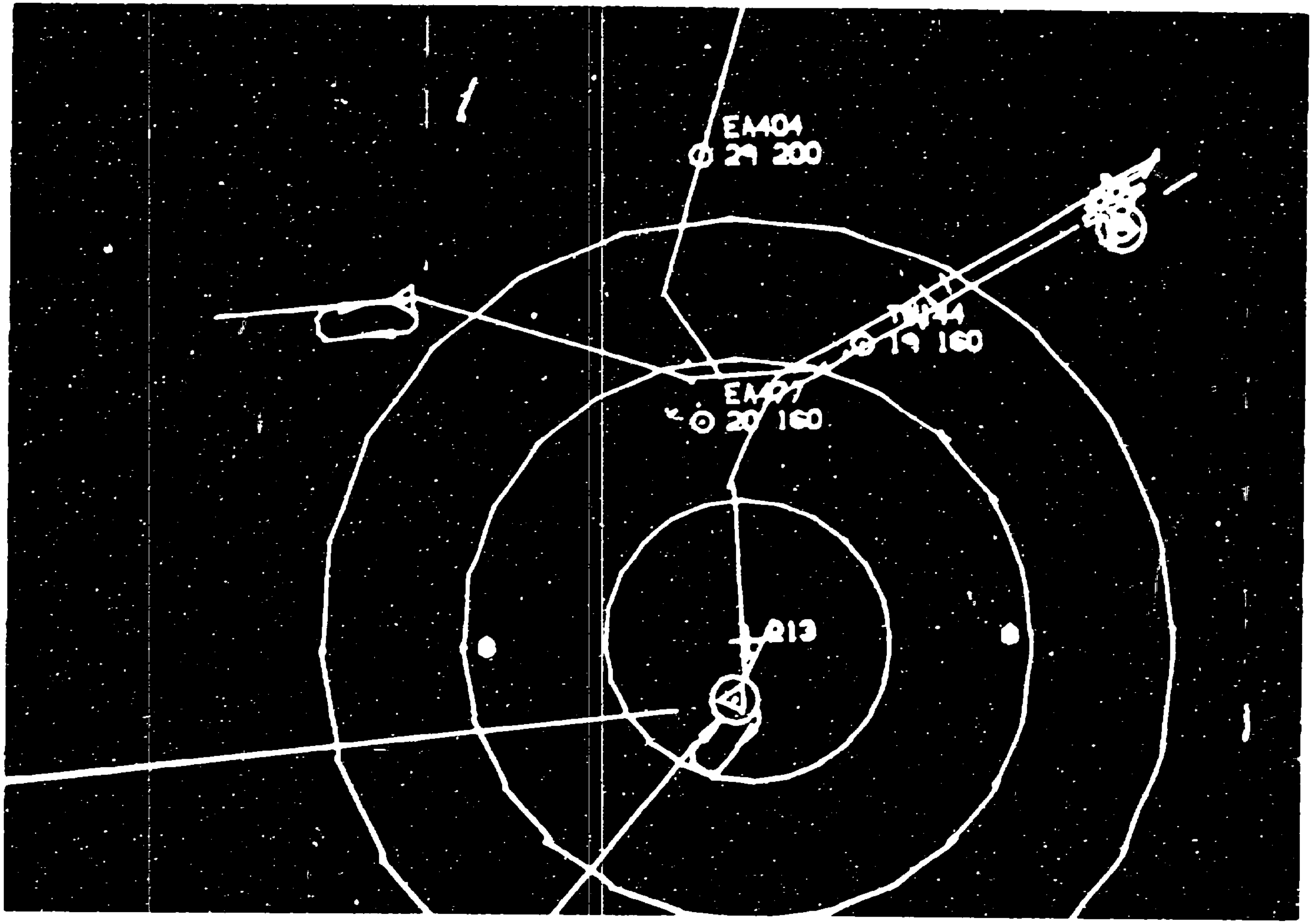
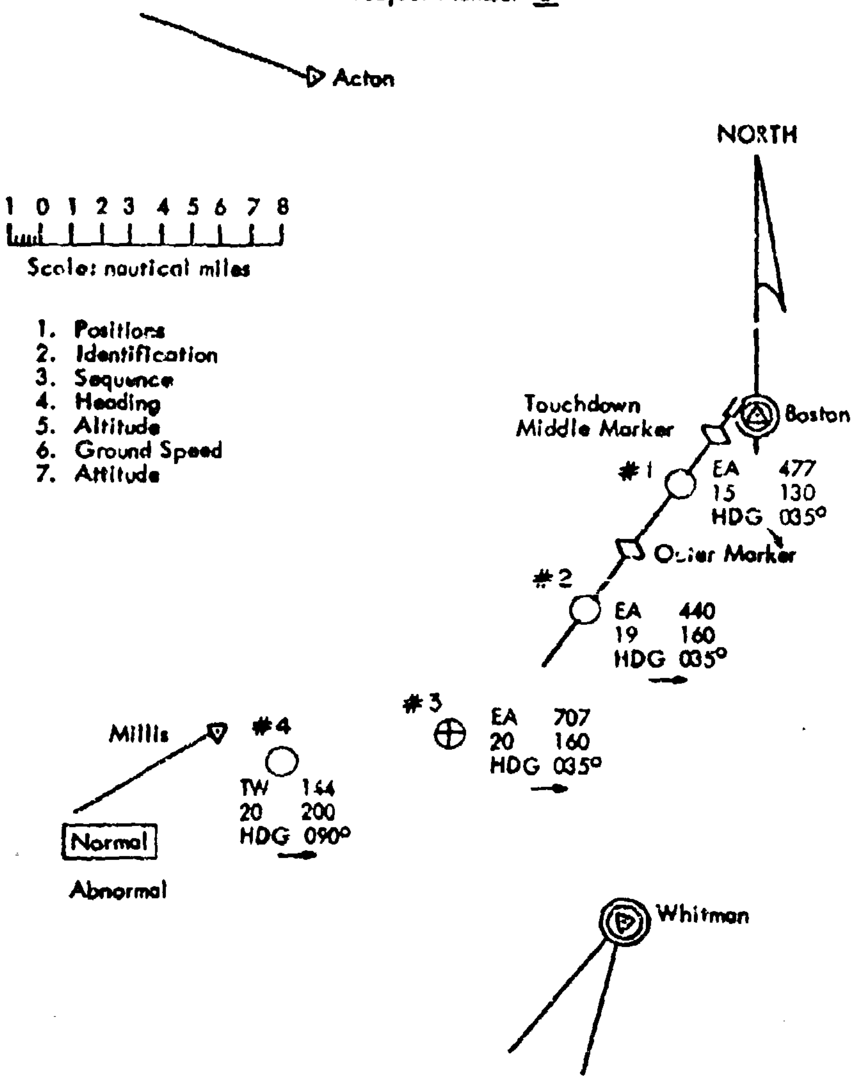


Figure 2.2.5 ATSD Presentation at Stop-action Quiz for Situation 4

TRAFFIC POSITION MAP

CASE FAA 1

Subject Number 2



Complete "Case Questionnaire" and announce that you are ready to continue

Fig. 2.2.6 Typical Completed Stop Action Quiz Sheet for Situation 1 could be graded on a right or wrong basis (i.e. identification, and landing sequence) were scored on a point system. Correct responses received a positive point score while incorrect responses received a negative point score. Position error was defined as the error in the pilot's estimate of the position of aircraft, including his own, with respect to the route structure. For all other information components, a component error was computed by subtracting the estimated value from the true value of the particular component in question. Thus both positive and negative errors were possible. Those cases where it was obvious that the subject's estimate for a given aircraft had it originating from the wrong holding fix, were scored as gross errors. Missing entries were recorded as null responses.

Spacing error was measured to determine the pilot's accuracy in estimating the other aircraft positions with respect to his own craft. It was computed in the same manner as the other information components.

2.3 Conflict Detection

Four of the seven traffic situations culminated with intrusions by other aircraft in the subject's airspace. Each intrusion was due to some abnormal event, and evidence of those events was provided to the pilot through either radio transmissions or the ATSD prior to the pause for the stop-action quiz. The subjects were required to specify on the quiz whether or not the traffic situation was normal. After the stop-action quiz, the simulation was continued until the point of closest approach (CPA) to the intruding aircraft had been reached. If the blunder had not been detected before the quiz and if the ATSD was not being used, there was little likelihood that the intrusion could be detected subsequent to the quiz and the simulations were not continued. Two of the conflict scenarios occurred during single runway approach situations, while the other two blunders occurred during independent operations on closely spaced parallel runways.

One of the single runway approach conflicts was the misinterpretation of a heading change instruction from the approach controller (situation 3). This resulted in a potential collision abeam of the subject's own aircraft which at the time was flying on the ILS. This blunder is depicted at a point of acute intrusion in Figure 2.3.1.

The second single runway conflict consisted of a radio blackout and subsequent failure to turn to a new heading, thus bringing the intruding aircraft into a head-on collision course with the subject's aircraft (situation 4). Figure 2.3.2 shows the ATSD at a point of acute intrusion for this situation.

The parallel runway conflicts were both essentially ILS crossover blunders. The first conflict had the intruding aircraft overshooting his ILS and acquiring the subject's ILS (situation 6). At the time of this blunder, the subject

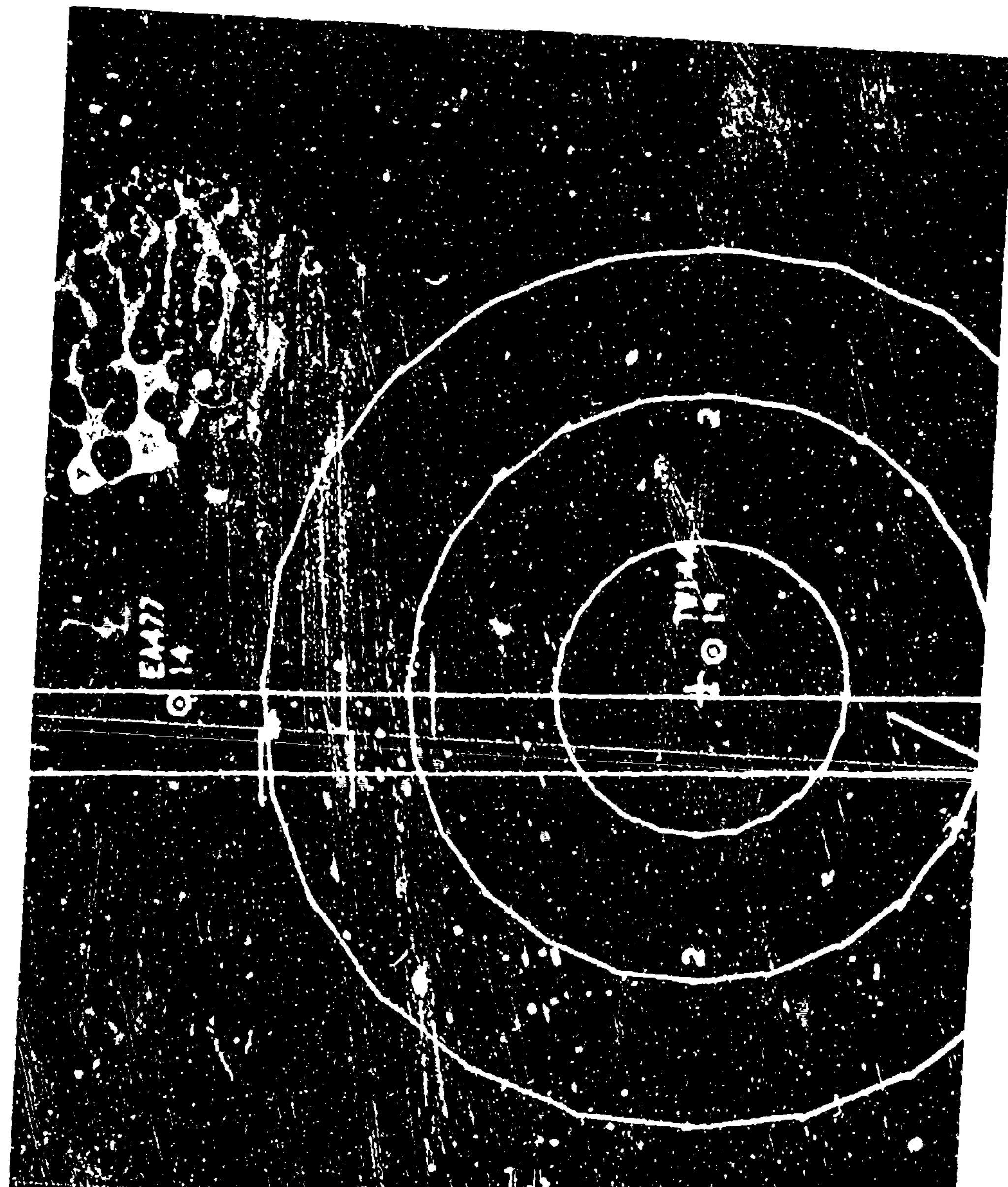


Figure 2.3.1 Situation 3 at Acute Intrusion

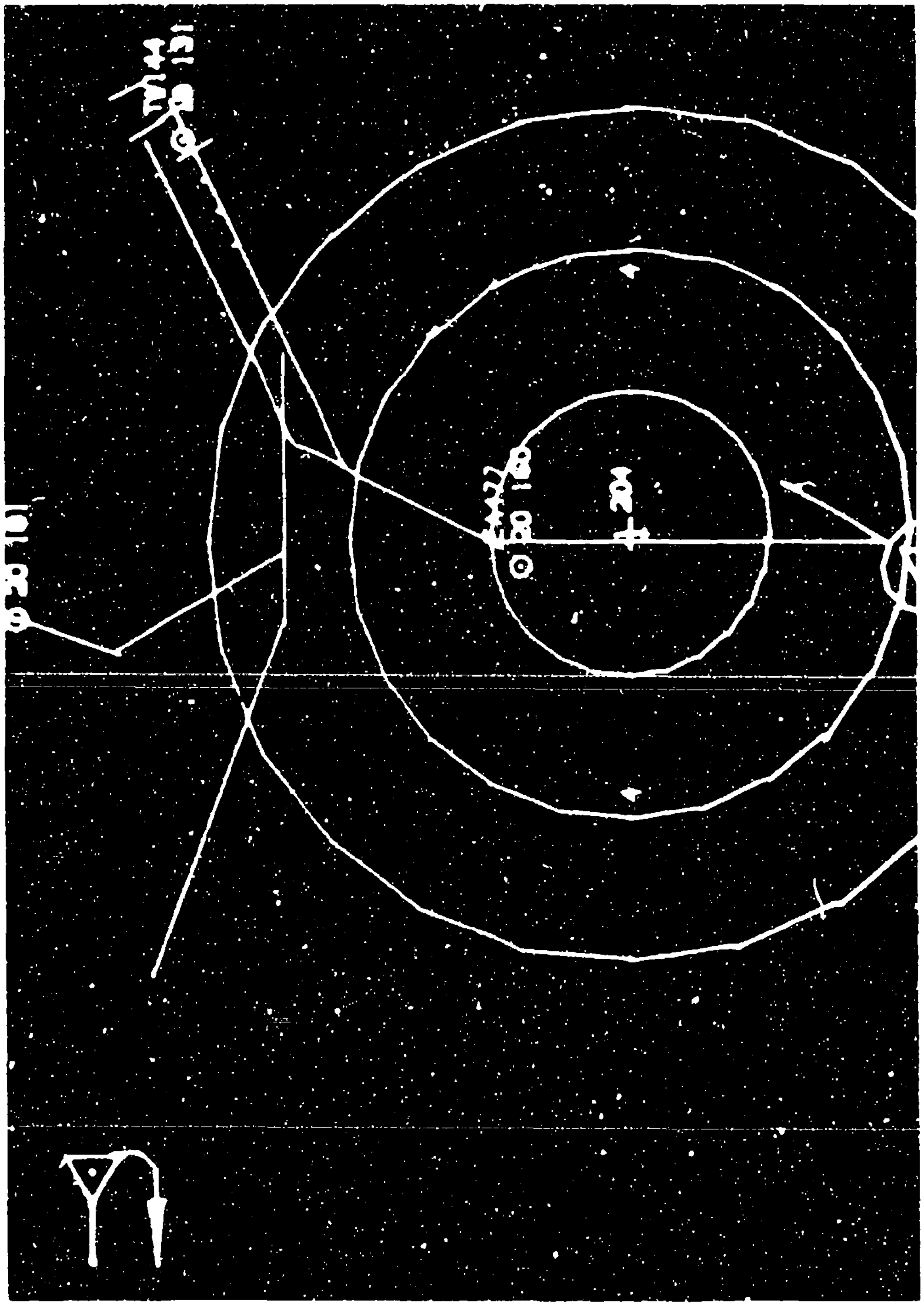


Figure 2.3.2 Situation 4 at Acute Intrusion

was in the process of acquiring his ILS. The second blunder had the intruding aircraft veering sharply (eighteen degree bank angle at 130 knots) from his ILS towards the subject aircraft after both planes had passed the outer marker (situation 7). At this time, the subject was busy flying the final approach glide path and preparing to land. The ATSD presentations at points of acute intrusion for these situations are depicted in Figures 2.3.3 and 2.3.4.

Pilot response times were measured by a strip chart recorder which monitored aileron deflection. Response time was measured from when the emergency alarm threshold was violated (whether or not the alarm actually sounded) until the beginning of the aileron deflection for the emergency escape manuever.

2.4 Alarms

In some cases, aural and visual alarms were used to warn the subjects of the close proximity of other aircraft or of an impending emergency.

The two mile proximity alert was an aural and visual alarm that was triggered whenever any aircraft violated a two mile separation criteria. It consisted of a low-keyed momentary aural signal accompanied by a blinking of the intruder's aircraft symbol on the ATSD. The ship symbol would continue blinking as long as the intruding aircraft remained within a two mile range. No altitude criteria was used for this alert because the experimental plan required

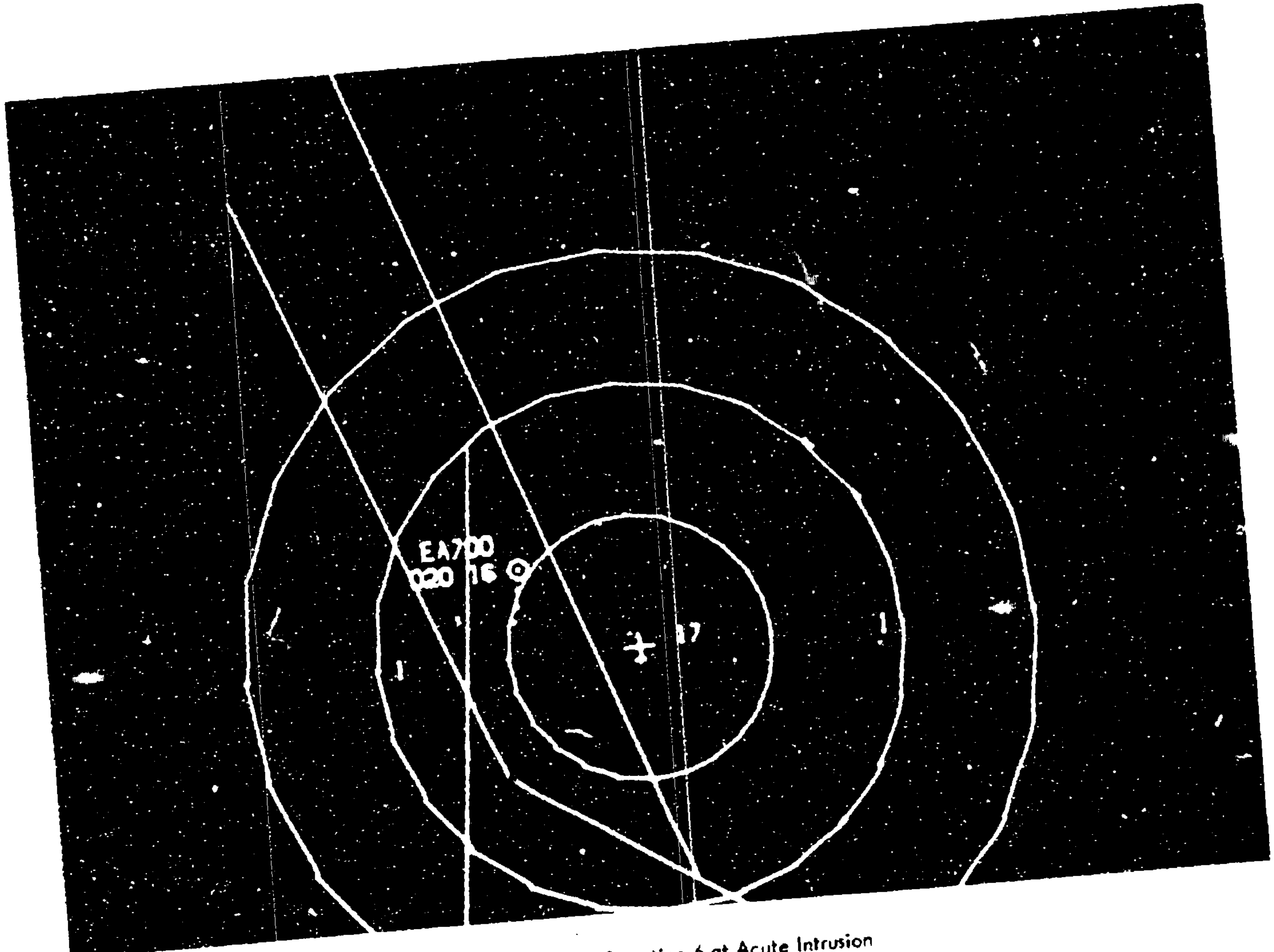


Figure 2.3.3 Situation 6 at Acute Intrusion

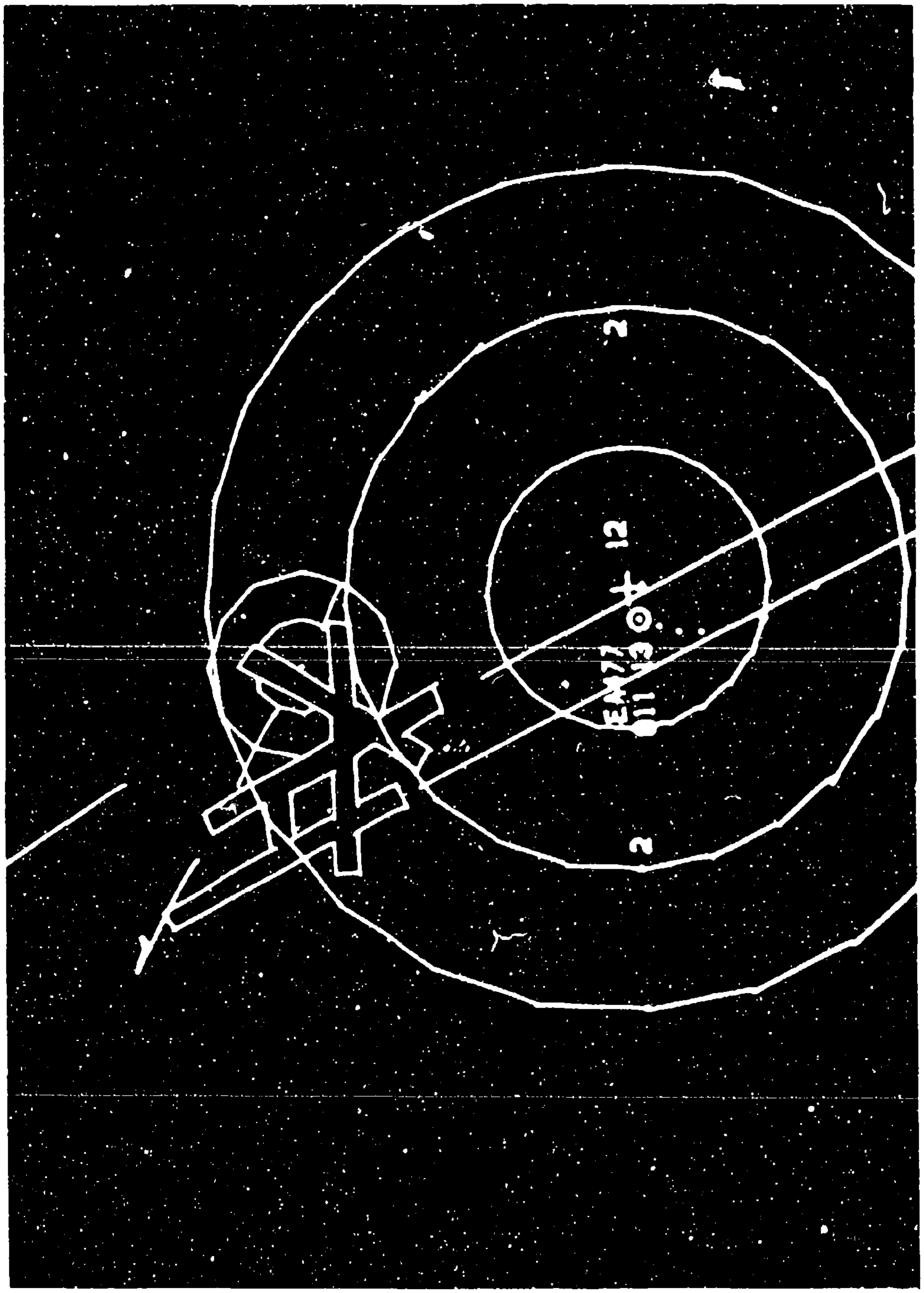


Figure 2.3.4 Situation 7 at Acute Intrusion

all subjects to react to the intrusion. In actual practice, some sort of altitude filtering would be implemented.

The violation of this alert criteria did not necessarily constitute a dire threat. It merely informed the pilot that there was another aircraft nearby. It was up to the pilot to evaluate the situation and take whatever action, if any, was required.

The emergency alarm was based on a Tau-like criteron and was used only during the parallel runway simulations. The signal consisted of an attention-demanding bell alarm. It was tripped whenever the nearest aircraft flying the adjacent ILS violated a Tau Criteria (lateral displacement divided by rate of change of lateral displacement with respect to the subject's ILS). The lateral displacement measurement was taken as the shortest distance from the intruder to the subject's ILS centerline. The rate of change measurement was taken as the rate of change of this distance. The Tau threshold was set at 28 seconds.

The simulations were designed such that the tripping of this alarm constituted a dire threat to the safety of the subject's aircraft. Upon hearing the alarm the subject was required to initiate a standard emergency manuever.

2.5 Emergency Procedure

A standard emergency procedure was designated for parallel runway operations. It consisted of the following two steps:

1) level off

2) perform a thirty degree banked turn at constant altitude away from the intruding aircraft.*

When the emergency alarm was employed, the subjects were required to initiate the manuever as soon as the alarm was tripped. In those cases where the alarm was not used, the subjects had to monitor the ATSD to detect when an overt threat had been established and then initiate the manuever. Detection of blunders causing significant track deviations was complicated by a natural ILS track wander and radar noise associated with the flight paths of the adjacent traffic elements.

The lateral constant-altitude manuever was chosen as a standard because of its ability to generate the greatest track deviation in the shortest period of time. It is realized that this manuever is not the best avoidance procedure under all circumstances, but within the context of the scenarios developed for these experiments, it provided the safest escape route. In addition, it provided an easily obtainable source of pilot reaction time measurements.

---

*A 30° bank does not exceed stall limits under these conditions, and it is believed that pilots will accept such a maneuver as an emergency procedure.

2.6 Subjects

All subjects used in both tests conducted by Howell and the results reported here were professional air carrier pilots, military pilots or commerical pilots with an instrument rating. Most of the participating pilots had an air transport rating.

Twenty pilots participated in the Phase I experiments and twenty-four pilots participated in Phase II. Of those pilots participating in the latter tests, twelve flew as two man flight crews (six crews), while the remaining twelve flew the simulations as single pilots. Five of the crews flew the single runway simulations and four of these crews also flew the parallel runway simulations. In addition, one other crew flew the parallel runway scenarios only. All the Phase I experiments were flown with single pilots. None of the Phase I subjects were used in Phase II.

2.7 Training

A three hour training session was used for both groups of pilots. Those pilots in the Phase I tests were considered adequately trained if they could close-up and follow another aircraft in trail at a specified distance with a standard deviation of aircraft separation not greater than one tenth nautical mile. If any pilot felt that he needed more training, that opportunity was made available to him.

Since the pilots participating in the Phase II experiments were not required to perform a spacing task, the previous training criteria was not appropriate. Instead, a nominal three hour training session was conducted. If, at the end of that session, both the experimenter and the subject felt that an adequate training level had been achieved, then training was terminated. If it was felt that performance was not adequate, then further training was provided.

Both training programs consisted of practice simulations that were representative of those that would be encountered during data gathering runs. Those pilots participating in the Phase II tests received more instruction in conflict detection and practice in taking the stop-action quiz than did Phase I pilots. Pilots flying as a two-man crew also received instruction in task allocation and coordination. Those pilots who would encounter the alarms in the test program were trained with these devices.

2.8 Flight Crew Task Allocation

During two man flight crew simulations, cockpit task responsibility was divided between the Captain and First Officer. For the most part, the division of duties corresponded to typical air carrier procedures. That is, the Captain was responsible for the inner loop control tasks of actually flying the aircraft, while the First Officer was responsible for handling communications, setting communication and navigation frequencies, cross checking the flight instruments and performing auxiliary control tasks, such as lowering the flaps and landing gear, at the Captain's request.

The procedures in this program differed from actual operations in that the ATSD had to be operated and monitored by the crew. The primary monitor of the ATSD was the First Officer. In addition, he would operate the range selection and information deletion controls. He also kept the Captain informed of pertinent traffic information (distance, altitude and ground speed of the nearest traffic elements). The Captain would monitor the ATSD as time permitted, but was generally engrossed in the inner loop control tasks.

2.9 Experimental Plan

2.9.1 Information Components

Information transfer aspects of pilot awareness were measured by pilot estimates of information components of the surrounding traffic elements as measured by the stop-action quiz. Stop-action quizzes were administered only during the four single runway situations.

Ten of the twelve single pilots and five of the crews participated in these tests. The four situations were presented in random order to these pilots. The independent variables in this experiment were crew treatment along with situation (scenario) treatment and alarm treatment. The dependent variables were the information component estimates.

2.9.2 Conflict Detection

Cognizance of a potential conflict on the single runway simulations was indicated by the appropriate response on the stop-action quiz or the pilot's radio transmission from the pilot to the controller. This data was acquired at the same time as the data on information components.

Awareness of a conflict during the parallel runway simulations was indicated by the performance of a standard emergency manuever. All twelve single pilots and five of the crews participated in these tests.

Of the single pilots participating, four had the aid of the alarms while eight did not. Of the eight single pilots who did not have the alarms, four had a one second information update rate while the remainder had a four second update rate. Two of the crews had the alarms while three did not. All crews had a four second update rate. The experimental master plan is shown in Table 2.9.1.

The independent experimental variables here were crew (one or two), alarm or no alarm, update period (1 or 4 seconds) and situation treatments. The dependent variables were detection and pilot reaction times (parallel simulations only).

TABLE 2.9.1

EXPERIMENTAL PLAN

CREW TREATMENT	SINGLE PILOTS										CREWS				
ALARM TREATMENT	ALARM				NO ALARM						ALARM		NO ALARM		
SUBJECT TREATMENT	2	4	6	10	1	3	5	7	8	9	21	24	22	23	25
ORDER OF PRESENTA-TION OF SITUATIONS	2 3 1 4	3 1 2 4	2 1 4 3	1 2 3 4	4 3 2 1	2 1 3 4	1 3 2 4	4 1 3 2	3 1 4 2	4 1 3 2	1 2 4 3	2 3 4 1	4 2 3 1	2 1 3 4	2 3 1 4

(a)

SINGLE RUNWAY SCENARIOS (SITUATIONS 1 thru 4)

CREW TREATMENT	SINGLE PILOTS												CREWS				
ALARM TREATMENT	ALARM				NO ALARM								ALARM		NO ALARM		
UPDATE RATE TREATMENT	FOUR SECOND				FOUR SECOND				ONE SECOND				FOUR SECOND		FOUR SECOND		
SUBJECT TREATMENT	2	4	6	10	1	3	5	11	7	8	9	12	21	26	22	23	25
ORDER OF PRESENTATION OF SITUATIONS	7 5 6	7 6 5	6 7 5	5 6 7	5 6 7	7 5 6	6 7 5	7 6 5	7 5 6	5 7 6	6 5 7	7 6 5	5 7 6	7 6 5	5 6 7	7 6 5	6 7 5

(b)

PARALLEL RUNWAY SCENARIOS (SITUATIONS 5 THRU 7)

CHAPTER 3

DATA SUMMARY AND ANALYSIS

The term "crew treatment" refers to the two sets of data selected to carry out the comparison between single pilots and two man crews. One data set represents various combinations of experimental factors which employed a single pilot, and the other set represents the same combination of experimental factors which employed two-man crews. The sets are balanced in every regard except for crew size. Similarly, groups of data were selected from the Phase I and Phase II data bases to carry out comparisons of the ATC treatment (no ATSD and party-line communications vs. ATSD and discrete address communications); of the spacing task treatment (with vs. without a spacing task); update period treatment (1 second vs. 4 seconds); and alarm treatment (alarm vs. no alarm). A "treatment", therefore, implies a specific grouping of data for the purpose of comparing experimental factors against each other by statistical analysis. A brief review of the statistical techniques employed in this report is present in the appendix.

In addition to testing for crew treatment effects, a primarly aim of this analysis is to incorporate the data taken during the Phase II experiments with portions of Howell's Phase I data. This combination of data is tested for the effect of two different air traffic control display/communications systems and the spacing task on pilot awareness. Howell's data concerning the present day air traffic control system (no ATSD and party-line communications) is compared to data from the Phase II tests taken in a discrete address communications environment with an ATSD. Neither set of tests involved the in-trail spacing task. Howell's tests with discrete address communications and the ATSD incorporated a spacing task. This data is compared to the data collected during the Phase II experiments using the same display-communication combination, but without the spacing task.

The analysis of these two major treatments employing two different data sources is based on two assumptions:

1. All subject pilots are representative of a homogeneous population of equal abilities and motivation.

2. Training for pilots from both data sources was comparable.

The statistical validity of the conclusions drawn from the analysis is, of course, correct only to the extent that the underlying assumptions are correct.

Since the data for the crew treatment was generated entirely by Phase II pilots, all of whom had the same training program, the above training assumption does not have to be employed. The homogeneity of the subject pilot population must, however, be assumed to validify the analysis.

The statistical model used to analyze all three major treatments (air traffic control systems, spacing task, and crew) was a fixed constants model with a variable number of replications.

This chapter concerns itself with the analysis of the data and a summary of statistical parameters. A discussion of the analytical results is included in the following chapters.

3.1 Analytical Design

3.1.1 Air Traffic Control Treatment

Only two of the four single runway scenarios were used in this analysis. One of these scenarios was a normal situation (situation 2) while the other was the heading read-back blunder (situation 3). No parallel runway results were compared since no equivalent data had been collected by Howell.

Stop-action quizzes were administered during both situations. Under the party-line communication channel with no ATSD condition, twelve subjects participated in the normal situation (2), while ten subjects participated in the blunder case (3). Under the discrete communication channel with ATSD condition, ten subjects participated in both the normal and blunder situations.

The air traffic control system treatment along with situation number and aircraft sequence number were the independent variables. The information component estimates and the number of conflict detections were the dependent variables.

3.1.2 Spacing Task Treatment

The same two single runway situations used to compare air traffic control treatments (Situations 2 and 3) were used to compare the spacing task treatments. In addition, the results from all three parallel runway simulations were included in the conflict detection evaluation.

3.1.2.1 Information Components

With the spacing task, seven pilots participated in the normal situation (2) while eight pilots participated in the blunder situation (3). Without the spacing task, ten pilots particpated in both situations. The stop-action quiz was administered during both scenarios.

The spacing task treatment as well as the situation and aircraft sequence were the independent variables while the information component estimates were the dependent variables.

3.1.2.2 Conflict Detection

As previously mentioned, eight and ten subjects flew the single runway blunder simulations with and without the spacing task respectively. All ten of the without-spacing-task subjects are included in the detection analysis at the stop-action quiz point, but only six are in the detection analysis at the closest point of approach since the others had alarms.

In the parallel runway scenarios, four pilots flew situation 5, seven pilots flew situation 6, and seven pilots flew situation 7 with the spacing task. Without the spacing task, twelve pilots flew in each of these cases. Of the twelve pilots not required to perform the spacing task, only eight (i.e. those not having the alarms) are considered in this analysis.

The independent variables here were the presence or absence of the spacing task, while the dependent variable was the detection (or non-detection) of conflicts.

3.1.3 Crew Treatment

The crew treatment analytical plan uses only Phase II data and follows the experimental plan presented in Table 2.9.1.

3.2 Pre-Analysis Data Summary

3.2.1 Information Components

A statistical summary (number of data points, mean, standard deviation, % gross errors, % null responses) for information component estimates of the data taken during the Phase II set of experiments are presented in Tables 3.2.1 thru 3.2.5. The data is pooled in accordance with the crew treatment. Position, spacing, altitude, ground speed and heading error information components are included. Summaries for the party-line/no ATSD system and for the discrete address/ATSD system with the spacing task are listed in the Phase I report.

The tables are broken down by situation and aircraft sequence within a given situation. Aircraft are indexed by their position relative to the subject in the landing sequence. The subject's aircraft is designated "0" while

TABLE 3.2.1

Crew Treatment: Position Error
nautical miles

	AIRCRAFT +2					AIRCRAFT +1					AIRCRAFT 0					AIRCRAFT -1				
	No.	M	S.D.	%N	%G	No.	M	S.D.	%N	%G	No.	M	S.D.	%N	%G	No.	M	S.D.	%N	%G
SITUATION 1	10	1.7	1.6	20	0	10	2.7	1.2	0	0	10	1.7	1.3	0	0	10	3.1	2.0	10	0
SITUATION 2	10	1.5	0.7	0	0	10	3.0	1.6	0	0	9	2.8	1.2	0	0	10	3.7	3.6	20	10
SITUATION 3	10	1.7	1.5	10	0	10	1.9	1.1	10	0	10	3.0	1.2	0	0	10	3.3	1.9	10	0
SITUATION 4	10	1.2	0.9	0	0	10	1.8	0.6	0	0	10	2.8	2.3	0	0	10	3.3	2.8	0	0

(a) SINGLE PILOTS

	AIRCRAFT +2					AIRCRAFT +1					AIRCRAFT 0					AIRCRAFT -1				
	No.	M	S.D.	%N	%G	No.	M	S.D	%N	%G	No.	M	S.D.	%N	%G	No.	M	S.D	%N	%G
SITUATION 1	5	1.2	0.7	0	0	5	1.6	1.4	0	0	5	0.9	0.6	0	0	5	2.7	1.5	0	0
SITUATION 2	4	1.7	1.3	0	0	4	2.6	2.5	0	0	4	2.7	1.1	0	0	4	2.4	2.8	0	0
SITUATION 3	5	1.4	1.0	0	0	5	1.7	1.2	0	0	5	2.0	1.3	0	0	5	3.1	1.9	0	0
SITUATION 4	5	1.4	0.6	0	0	5	2.1	0.6	0	0	5	1.5	0.7	0	0	5	4.7	2.9	20	0

(b) CREWS

No. = data samples
M = Mean
SD = standard deviation
%N = percent null answers
%G = percent gross errors (wrong sequence or fix)

Aircraft +2 - two slots ahead of ownship
Aircraft +1 - one slot ahead of ownship
Aircraft 0 - ownship
Aircraft -1 - one slot behind ownship

TABLE 3.2.2

Crew Treatment: Spacing Error
nautical miles

	AIRCRAFT +2				AIRCRAFT +1				AIRCRAFT -1			
	No.	M	S.D.	%N	No.	M	S.D.	%N	No.	M	S.D.	%N
SITUATION 1	10	1.5	1.8	20	10	1.4	1.0	0	10	-2.6	1.5	10
SITUATION 2	9	0.7	1.9	0	9	-0.4	2.6	0	9	1.3	2.0	22
SITUATION 3	10	-0.7	2.6	10	10	-0.3	1.7	10	10	-2.7	2.7	10
SITUATION 4	10	0.8	1.8	0	10	1.4	1.7	0	10	4.0	3.3	0

(a) SINGLE PILOTS

	AIRCRAFT +2				AIRCRAFT +1				AIRCRAFT -1			
	No.	M	S.D.	%N	No.	M	S.D.	%N	No.	M	S.D.	%N
SITUATION 1	5	0.0	0.6	0	5	1.0	0.7	0	5	-2.5	1.4	0
SITUATION 2	4	-0.4	2.9	0	4	0.7	2.2	0	4	1.3	2.8	0
SITUATION 3	5	-1.6	0.6	0	5	-1.0	0.7	0	5	-2.8	3.0	0
SITUATION 4	5	0.1	1.8	0	5	1.2	1.5	0	4	4.9	4.4	20

(b) CREWS

TABLE 3.2.3

Crew Treatment: Altitude Error
hundreds of feet

	AIRCRAFT +2				AIRCRAFT +1				AIRCRAFT -1			
	No.	M	S.D.	%N	No.	M	S.D.	%N	No.	M	S.D.	%N
SITUATION 1	10	-4.0	5.0	50	10	0.3	3.1	10	10	0.1	0.4	30
SITUATION 2	10	0.1	1.8	10	10	-5.0	10.7	20	10	5.6	11.3	10
SITUATION 3	10	-2.5	2.4	40	10	-.9	1.3	10	10	-.1	.3	20
SITUATION 4	10	-.1	1.2	0	10	-.8	3.3	0	10	-5.5	13.5	40

(a) SINGLE PILOTS

	AIRCRAFT + 2				AIRCRAFT + 1				AIRCRAFT - 1			
	No.	M	S.D.	%N	No.	M	S.D.	%N	No.	M	S.D.	%N
SITUATION 1	5	1.5	4.3	20	5	-.6	1.3	0	5	0	0	0
SITUATION 2	4	1.7	2.5	0	4	0	0	0	4	0	0	25
SITUATION 3	5	0.8	5.9	0	5	-.4	0.5	0	5	-.4	0.5	0
SITUATION 4	5	-.4	.9	0	5	0	0	0	5	-6.8	12.9	0

(b) CREWS

TABLE 3.2.4

Crew Treatment: Ground Speed Error
knots

	AIRCRAFT +2				AIRCRAFT +1				AIRCRAFT -1			
	No.	M	S.D.	%N	No.	M	S.D.	%N	No.	M	S.D.	%N
SITUATION 1	10	1	0	50	10	-17	15	20	10	-16	19	30
SITUATION 2	10	10	16	20	10	16	20	10	10	0	0	10
SITUATION 3	10	-12	17	20	10	-5	16	0	10	16	33	30
SITUATION 4	10	-7	15	0	10	-7	14	10	10	0	0	40

(a) SINGLE PILOTS

	AIRCRAFT +2				AIRCRAFT +1				AIRCRAFT -1			
	No.	M	S.D.	%N	No.	M	S.D.	%N	No.	M	S.D.	%N
SITUATION 1	5	0.7	0.5	10	5	0.6	0.5	0	5	-5.4	17.6	0
SITUATION 2	4	9.7	14.2	0	4	9.3	20.6	0	4	0	0	0
SITUATION 3	5	-4.4	11.3	0	5	-6.0	19.2	0	5	7.5	9.9	20
SITUATION 4	5	5.2	11.4	0	5	-6.0	19.2	0	5	0	0	0

(b) CREWS

TABLE 3.2.5

Crew Treatment: Heading Error
degrees

	AIRCRAFT +2				AIRCRAFT +1				AIRCRAFT -1			
	No.	M	S.D.	%N	No.	M	S.D.	%N	No.	M	S.D.	%N
SITUATION 1	10	-2	3	40	10	-1	3	30	10	6	14	30
SITUATION 2	10	-1	4	40	10	-2	33	40	10	-4	9	40
SITUATION 3	10	-1	3	30	10	-1	3	30	10	113	25	40
SITUATION 4	10	0	3	30	10	43	61	30	10	-2	6	40

(a) SINGLE PILOTS

	AIRCRAFT +2				AIRCRAFT +1				AIRCRAFT -1			
	No.	M	S.D.	%N	No.	M	S.D.	%N	No.	M	S.D.	%N
SITUATION 1	5	0	0	40	5	0	0	20	5	-5.0	18.0	20
SITUATION 2	4	0	0	50	4	-6.7	11.3	25	4	-4.0	5.2	25
SITUATION 3	5	0	0	20	5	0	0	20	5	122.5	23	20
SITUATION 4	5	0	0	40	5	77.5	57.2	20	5	15	17.7	40

(b) CREWS

the aircraft just ahead and behind in the sequence are designated "+1" and "-1" respectively. The aircraft two slots ahead and behind are designated "+2" and "-2" respectively, and so forth. Most simulations included three or four aircraft besides the subject's. Only the -1 thru +2 aircraft are included in the analysis since these are the only aircraft to appear in all simulations.

Composite graphs depicting the absolute value of the mean information component errors are presented in Figures 3.2.1 thru 3.2.17.

In both the tables and the graphs the following heading and label definitions are used.

TODAY'S SYSTEM	=	No ATSD and party-line communications (Howell's data)
FUTURE SYSTEM	=	ATSD, discrete communications, no spacing task (Phase II data: single pilots only)
SPACING TASK	=	ATSD, discrete communications, spacing task (Howell's data)
NO SPACING TASK	=	ATSD, discrete communications, no spacing task (Phase II data: single pilots only)
SINGLE PILOTS	=	ATSD, discrete communications, no spacing task (Phase II data: single pilots only)
CREW	=	ATSD, discrete communications, no spacing task (Phase II data: two man crews only)

3.2.2 Conflict Detection

The reaction time statistics under the crew treatment during the parallel runway simulations are summarized in

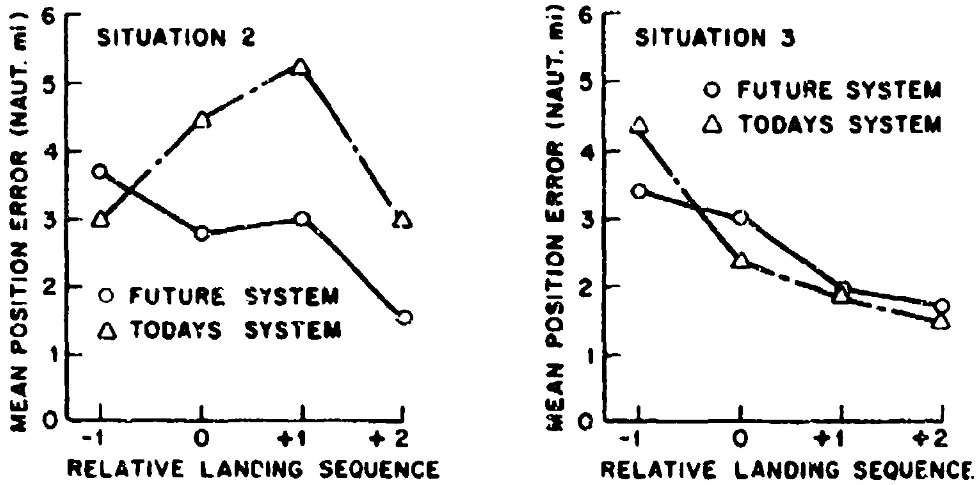


Figure 3.2.1

ATC TREATMENT: COMPOSITE GRAPHS OF POSITION ERROR
(PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

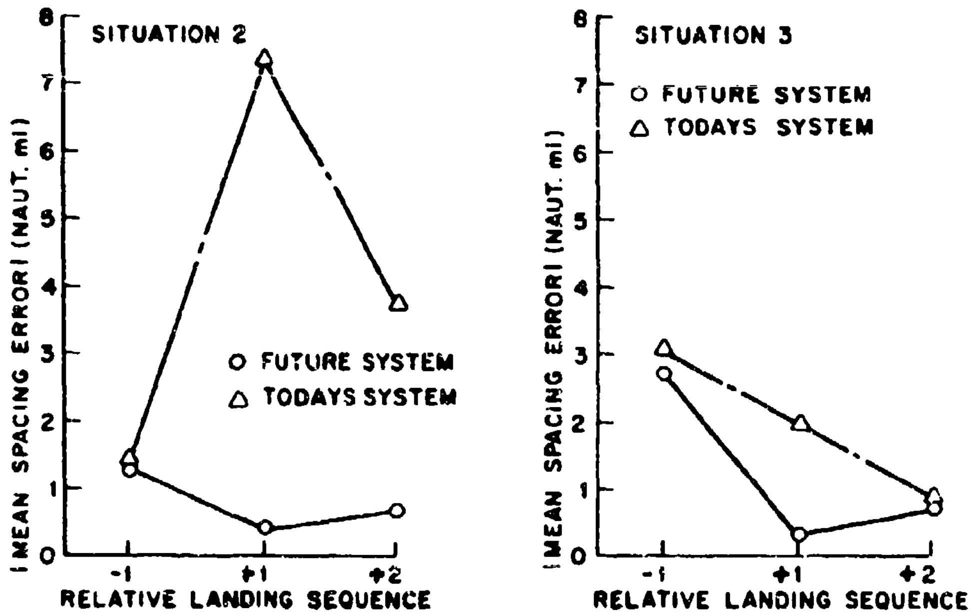


Figure 3.2.2

ATC TREATMENT: COMPOSITE GRAPHS OF SPACING ERROR
(PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

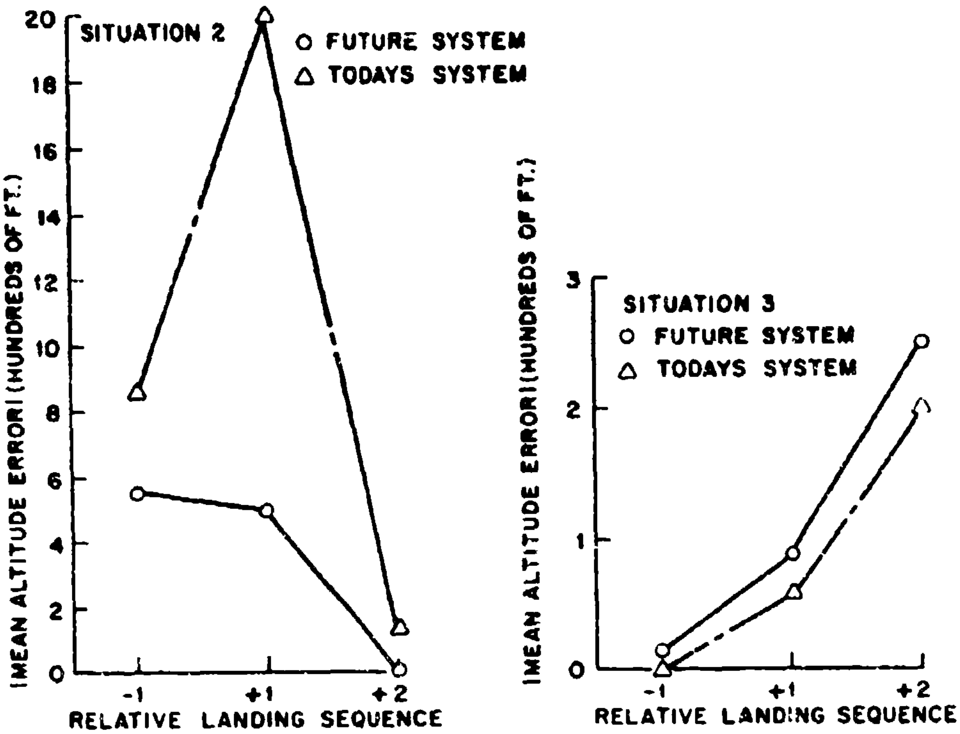


Figure 3.2.3

ATC TREATMENT: COMPOSITE GRAPHS OF ALTITUDE ERROR
(PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

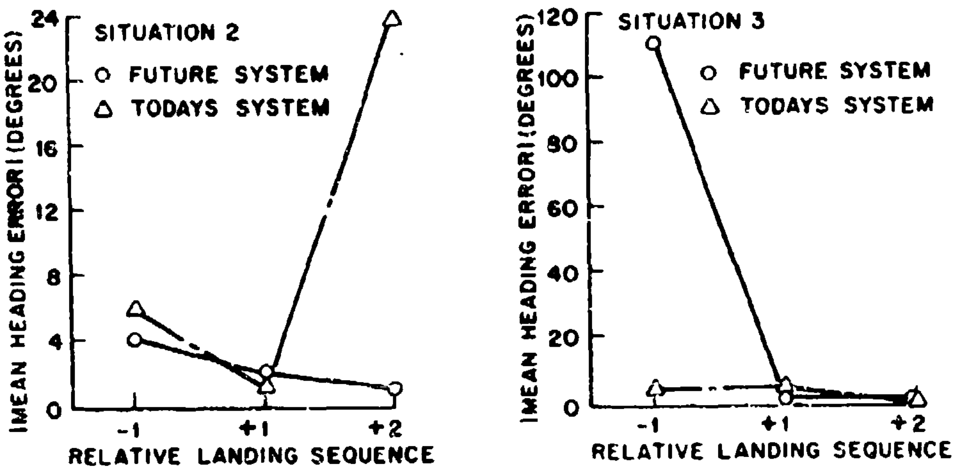


Figure 3.2.4

ATC TREATMENT: COMPOSITE GRAPHS OF HEADING ERROR
(PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

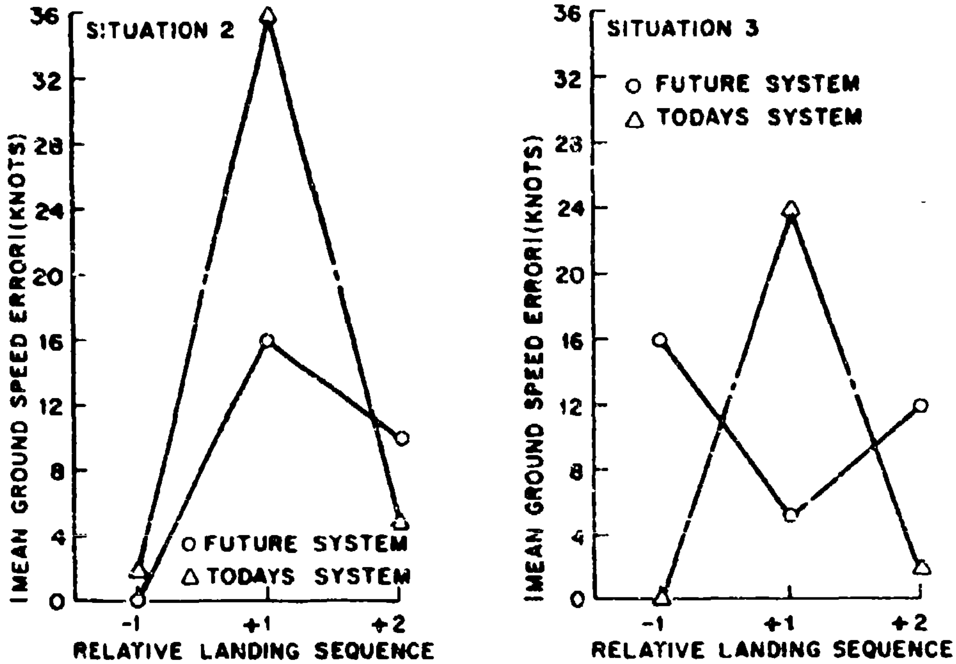


Figure 3.2.5

ATC TREATMENT: COMPOSITE GRAPHS OF GROUND SPEED ERROR
(PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

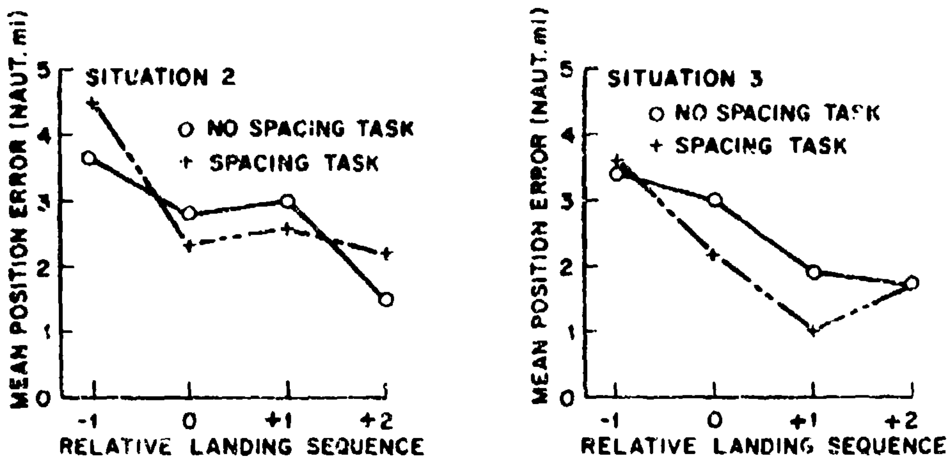


Figure 3.2.6

SPACING TASK TREATMENT: COMPOSITE GRAPHS OF POSITION ERROR (PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

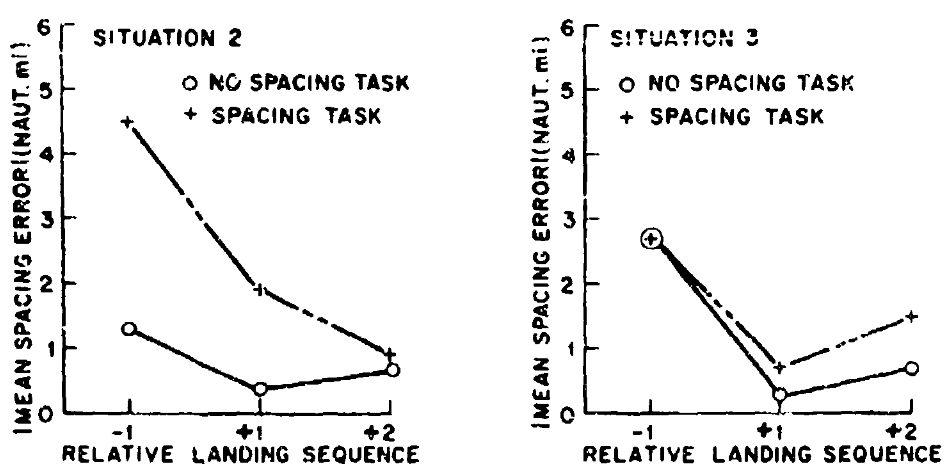


Figure 3.2.7

SPACING TASK TREATMENT: COMPOSITE GRAPHS OF SPACING ERROR (PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

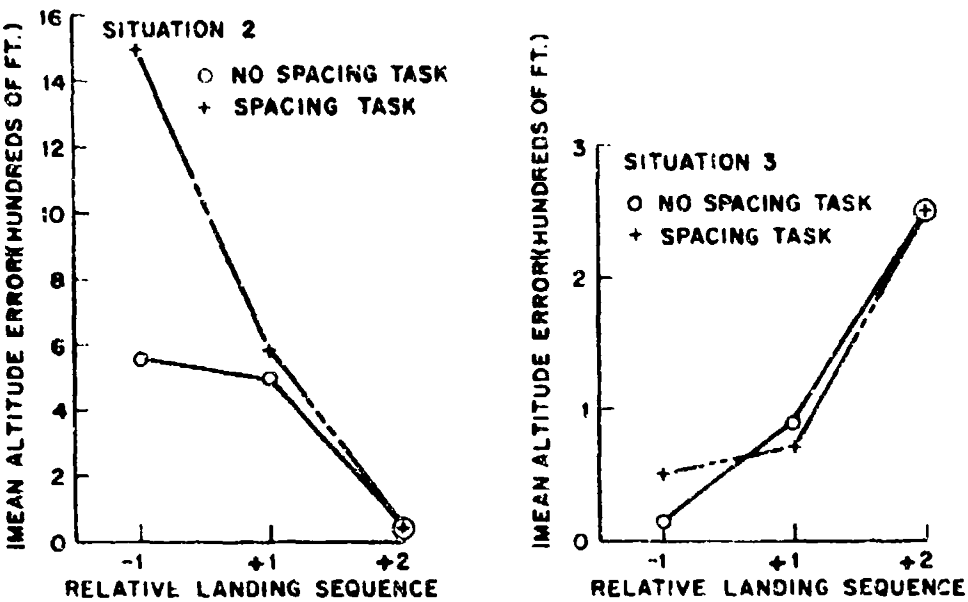


Figure 3.2.8

SPACING TASK TREATMENT: COMPOSITE GRAPHS OF ALTITUDE ERROR
(PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

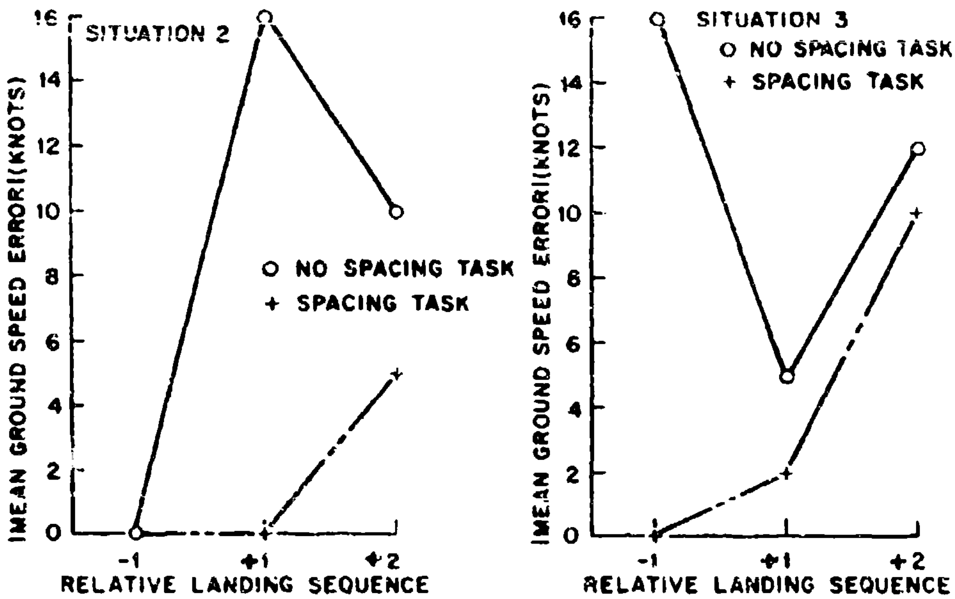


Figure 3.2.9

SPACING TASK TREATMENT: COMPOSITE GRAPHS OF GROUND SPEED ERROR
(PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

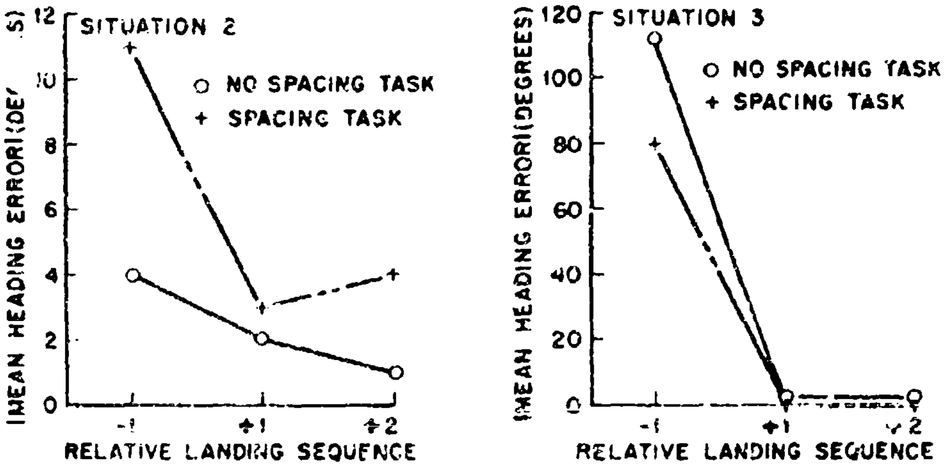


Figure 3.2.10

SPACING TASK TREATMENT: COMPOSITE GRAPHS OF HEADING ERROR
(PRE-ANALYSIS: PLOTTED SITUATION AND AIRCRAFT)

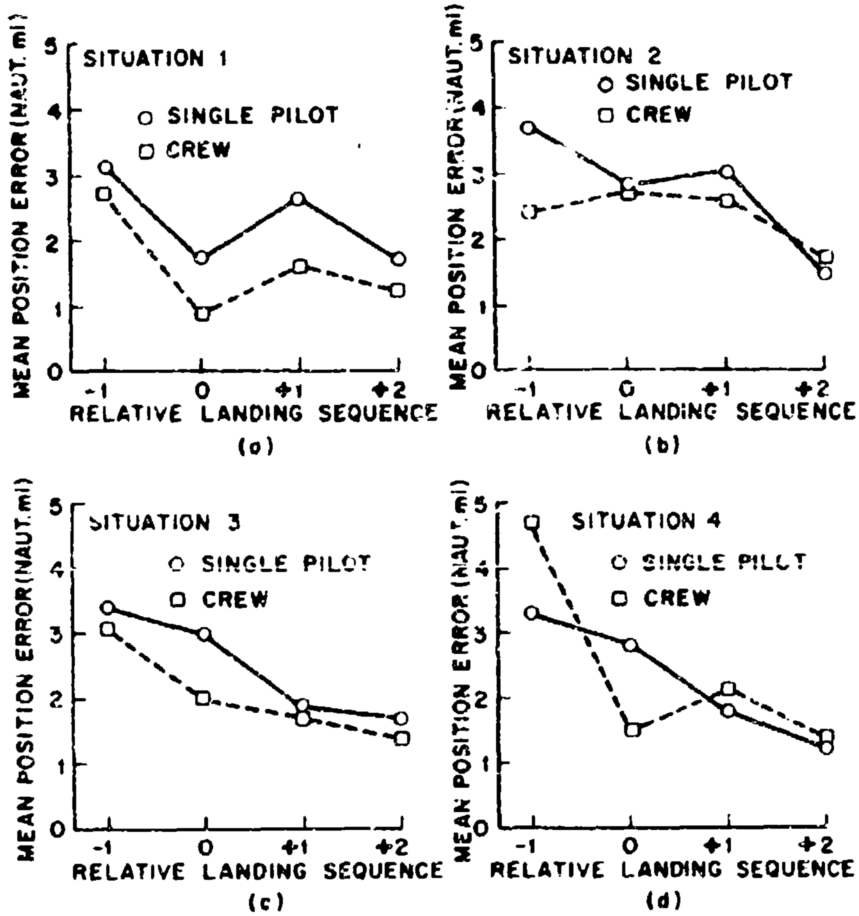


Figure 3.2.11

CREW TREATMENT: COMPOSITE GRAPHS OF POSITION ERROR
(PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

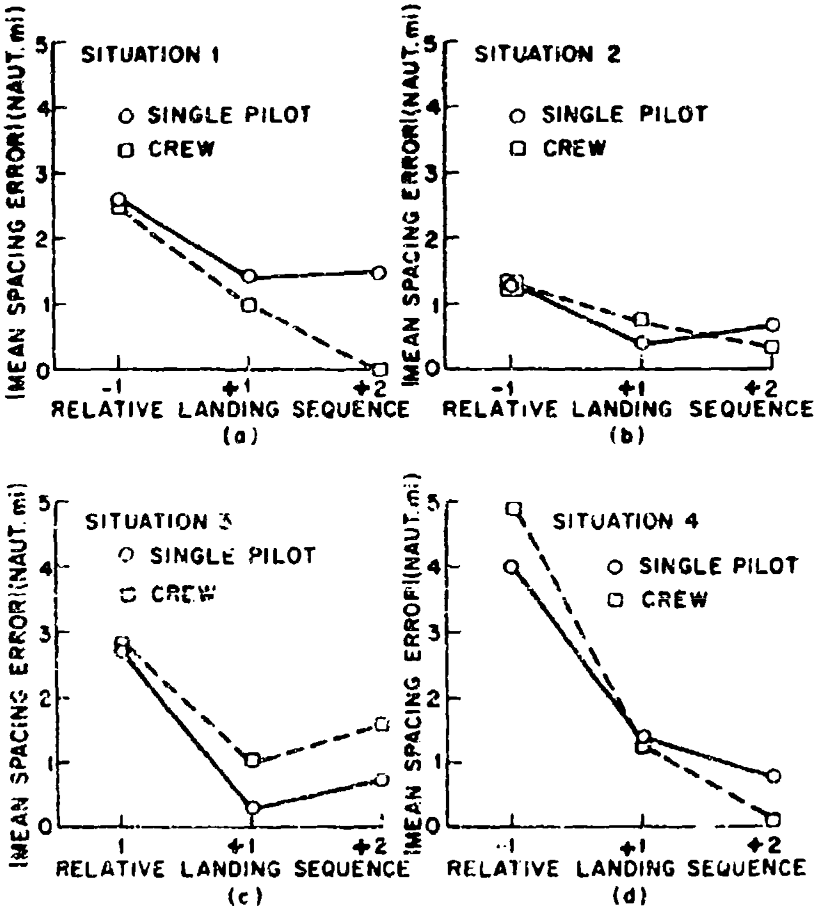


Figure 3.2.12

CREW TREATMENT: COMPOSITE GRAPHS OF SPACING ERROR (PRE-ANAYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

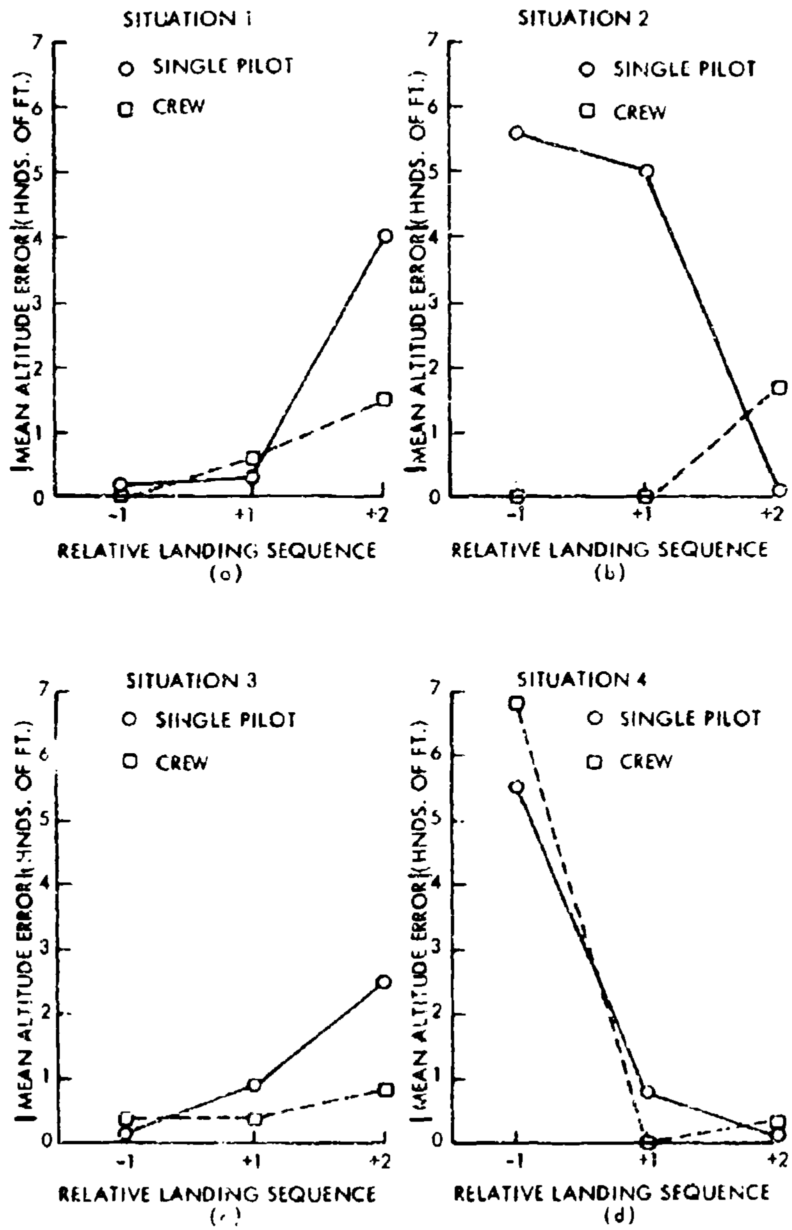


Figure 3.2.13

CREW TREATMENT: COMPOSITE GRAPHS OF ALTITUDE ERROR (PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

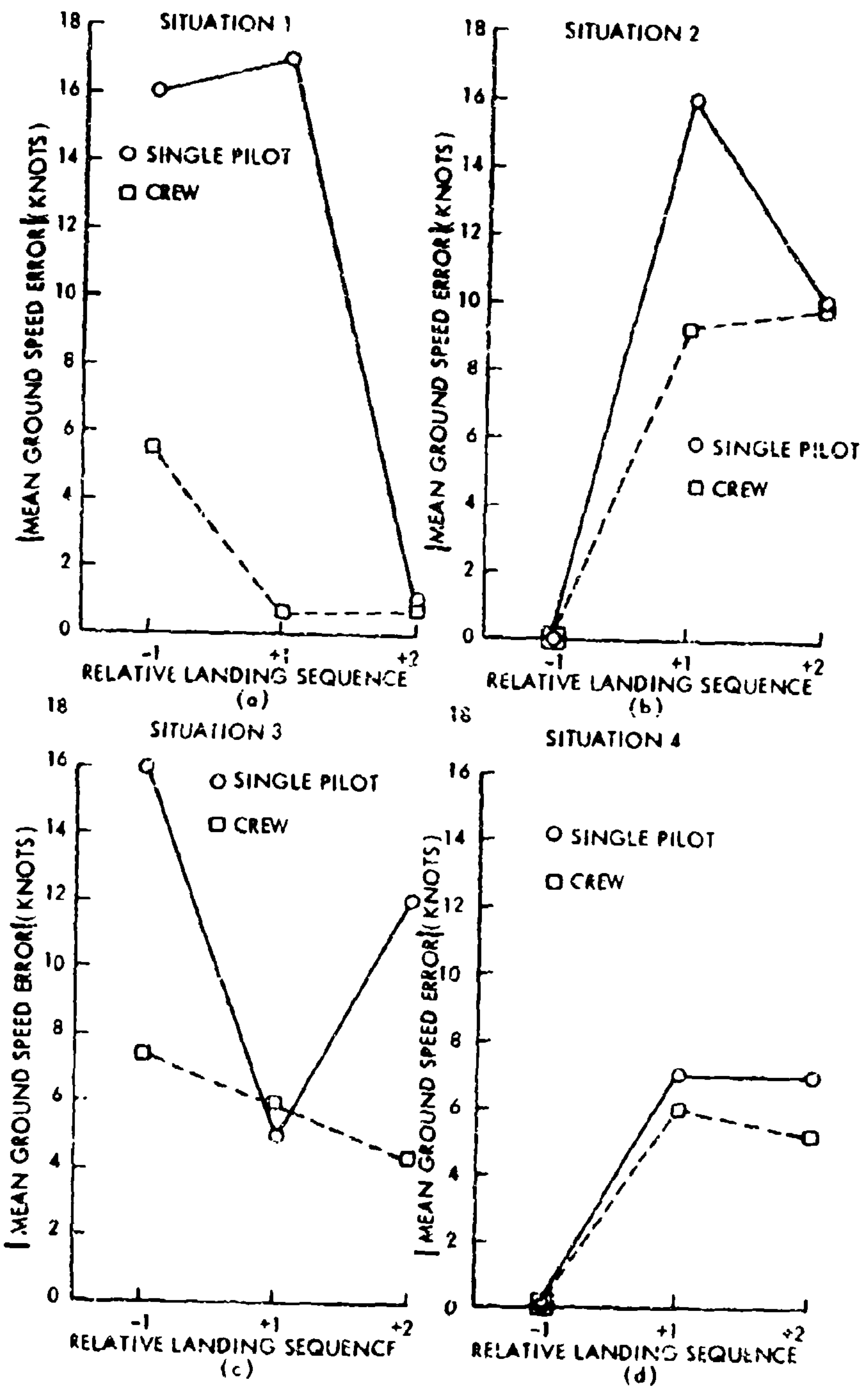


Figure 3.2.14

CREW TREATMENT: COMPOSITE GRAPHS OF GROUND SPEED ERROR (PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

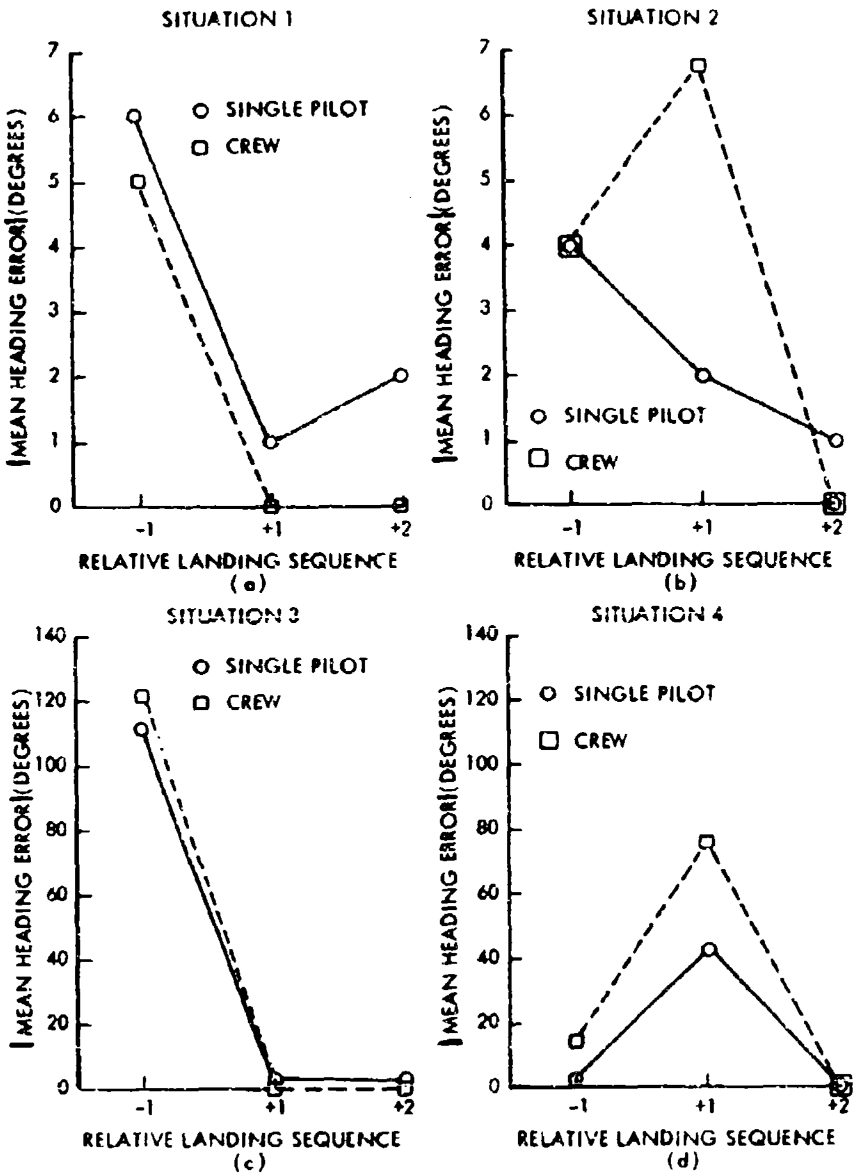


Figure 3.2.15

CREW TREATMENT: COMPOSITE GRAPHS OF HEADING ERROR
(PRE-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

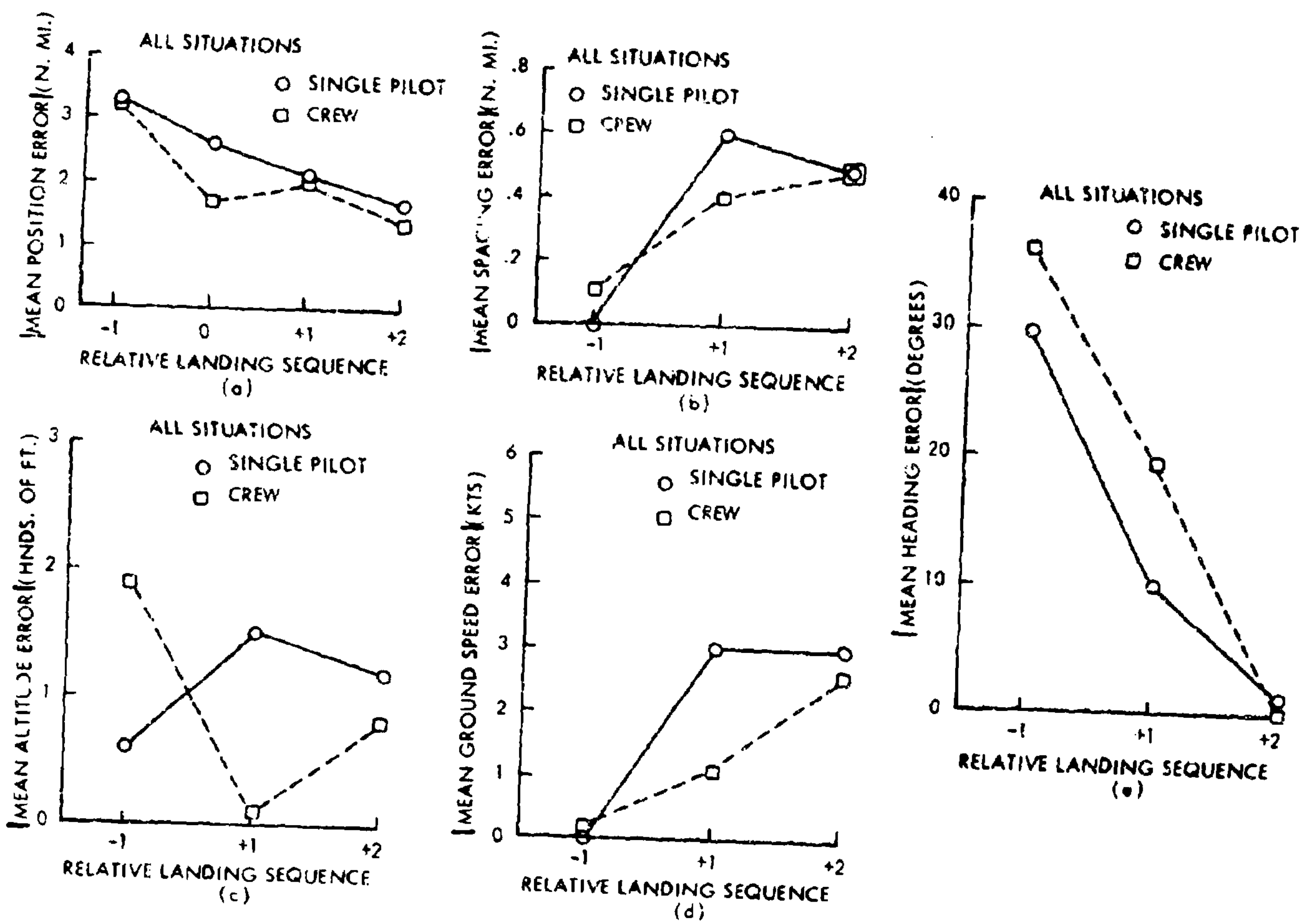


Figure 3.2.16 CREW TREATMENT: COMPOSITE GRAPHS OF INFORMATION ERROR COMPONENTS (PRE-ANALYSIS: PLOTTED BY AIRCRAFT)

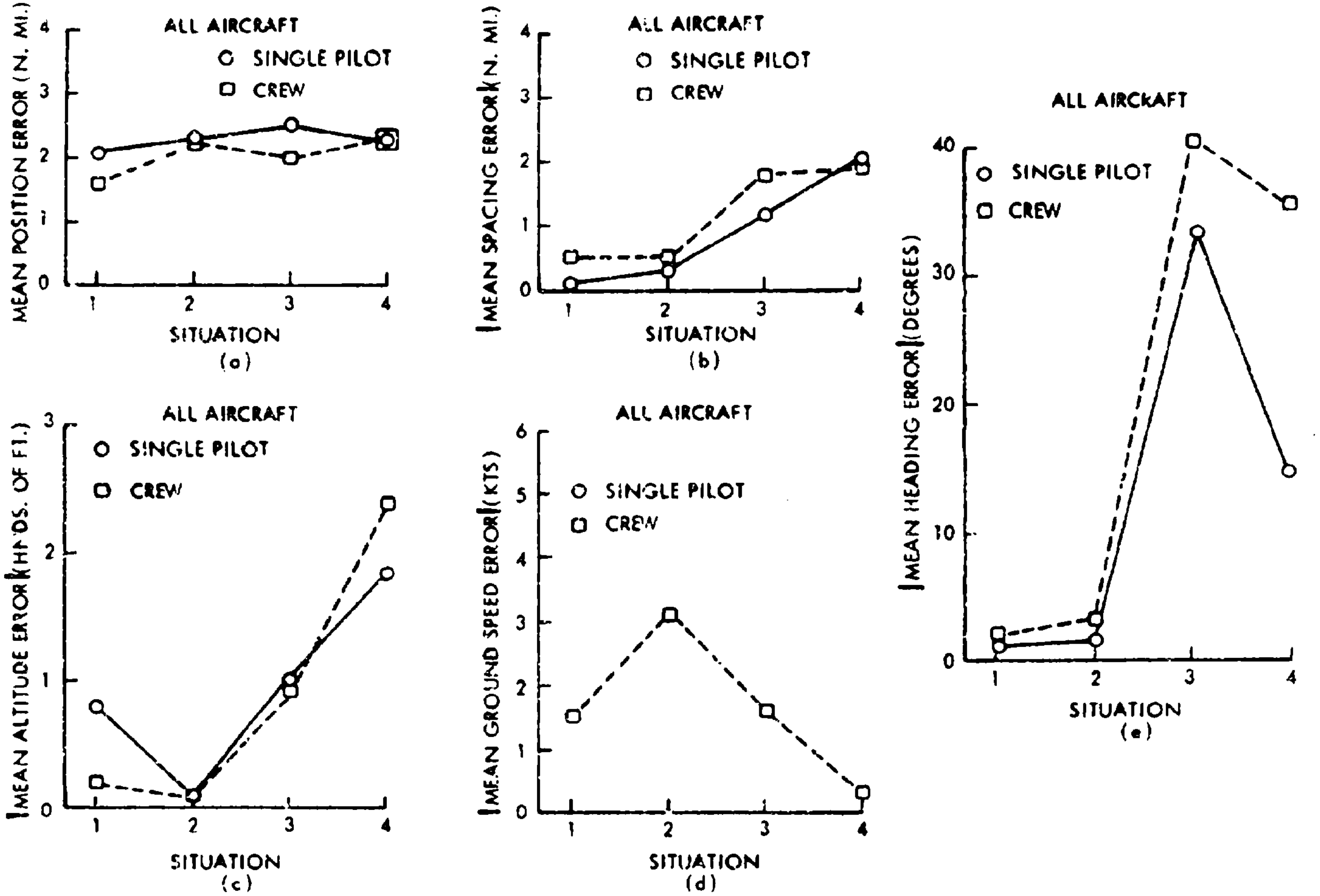


Figure 3.2.17 CREW TREATMENT: COMPOSITE GRAPHS OF INFORMATION ERROR COMPONENTS (PRE-ANALYSIS: PLOTTED BY SITUATION)

Table 3.2.6. Reaction time was defined as the time delay between the beginning of the intruding aircraft's blunder and the initiation of the subject's lateral escape manuever.

3.3 Analysis*

3.3.1 Information Components

Throughout all of the analysis in this chapter, it is assumed that the replications in any given treatment represent independent estimates of the parameter being measured.

A three way factorial analysis of variance[9] was performed on each information error component for all major treatment variables. The main experimental effects tested were the major treatment variables (air traffic control, spacing task or crew treatment), situation treatment, and aircraft sequence treatment (+2, +1, etc.). First and second order interactions were tested. If an interaction involving only two main effects was found significant, then a first level (by situation) breakdown analysis of variance was performed. If significant interactions involving all three main effects during the original analysis were encountered or a significant interaction involving two of the main effects during a first level breakdown analysis were encountered, then a second level breakdown (by situation and aircraft) analysis of variance was performed.

---

*See Appendix for explanation of statistical methods.

TABLE 3.2.6

Reaction Time Statistical Summary
seconds

CREW TREATMENT	SINGLE PILOTS									CREWS					
ALARM TREATMENT	ALARM			NO ALARM						ALARM			NO ALARM		
UPDATE RATE TREATMENT	FOUR SECONDS			FOUR SECONDS			ONE SECOND			FOUR SECONDS			FOUR SECONDS		
STATISTICAL PARAMETER	No.	M	S.D.	No.	M	S.D.	No.	M	S.D.	No.	M	S.D.	No.	M	S.D.
SITUATION 6	3	5.1	0.3	4	11.7	3.5	4	10.9	2.4	1	5.7	∞	2	13.4	4.9
SITUATION 7	4	10.8	0.3	4	10.8	2.4	4	12.1	2.5	2	13.1	3.3	3	13.5	3.2

No. ■ Number of Data Points
M ■ Mean
S.D. ■ Standard Deviation

The results of the original and first level breakdown analyses are summarized in Tables 3.3.1 thru 3.3.5. The results of the second level breakdown analyses are incorporated in Tables 3.3.6 thru 3.3.8 which indicate statistical significance (or nonsignificance) of the main effects.

Statistical signficance was achieved if the F ratio bettered the five percent level ($p<.05$). In calculating F, the effect being tested was always used as the numerator while the residual term was used as the denominator. This accounts for some of the values of F being less than one.

Contingency tables were used to judge signficance for both the information point scores (identification and landing sequence) and the null responses. No identification component analysis was performed for the air traffic control or spacing task treatment because data in an appropriate format was unavailable. In both cases, a Chi Square test was performed to test a situation and an aircraft null hypothesis[8]. The hypotheses tested were the following:

Situation Hypothesis: There is no difference in point scores (or null responses) due to situations (all aircraft pooled).

Aircraft Hypothesis: There is no difference in point scores (or null responses) due to aircraft sequence (all situations pooled).

If, for a given major treatment variable, these hypotheses were verified then all situations and aircraft were pooled in constructing the contingency tables. If either or both

TABLE 3.3.1

ATC TREATMENT: RESULTS OF THREE WAY ANALYSIS OF VARIANCE FOR INFORMATION COMPONENTS

Information Component		Main Effects			First Order Interactions			Second Order Interactions
		ATC	SITUATION	AIRCRAFT SEQUENCE	ATC x SITUATION	ATC x AIRCRAFT SEQUENCE	SITUATION x AIRCRAFT SEQUENCE	ATC x SITUATION x AIRCRAFT SEQUENCE
Position	F D.F.	4.5 1:134 SIG.	5.4 1:134 SIG.	3.6 3:134 N.S.	2.5 1:134 N.S.	0.3 3:134 N.S.	2.5 3:134 N.S.	2.7 3:134 SIG.
Spacing	F D.F.	18.8 1:37 SIG.	0.8 1:97 N.S.	1.7 2:97 N.S.	8.7 1:97 SIG.	2.1 2:97 N.S.	5.9 2:97 SIG.	=0 2:95 N.S.
Altitude	F D.F.	1.7 1:82 N.S.	0.6 1:62 N.S.	12.5 2:82 SIG.	3.4 1:82 N.S.	3.2 2:82 SIG.	11.5 2:82 SIG.	2.7 2:80 N.S.
Ground Speed	F D.F.	15:5 1:77 SIG	0.6 1:77 N.S.	4.6 2:77 SIG.	4.3 1:77 SIG.	8.9 2:77 SIG.	2.0 2:77 N.S.	4.7 2:77 SIG.
Heading	F D.F.	3.7 1:74 N.S.	2.1 1:74 N.S.	4.7 2:74 SIG.	10.6 1:74 SIG.	9.7 2:74 SIG.	11.3 2:74 SIG.	0.7 2:74 N.S.

F ≡ F RATIO

DF ≡ DEGREES OF FREEDOM

N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)

SIG. ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

TABLE 3.3.2

SPACING TASK TREATMENT: RESULTS OF THREE WAY ANALYSIS OF VARIANCE FOR INFORMATION COMPONENTS

Information Components		Main Effects			First Order Interactions			Second Order Interactions
		SPACING TASK	SITUATION	AIRCRAFT SEQUENCE	SPACING X SITUATION	SPACING X AIRCRAFT SEQUENCE	SITUATION X AIRCRAFT SEQUENCE	SPACING X SITUATION X AIRCRAFT SEQUENCE
POSITION	F D.F.	0.5 1:116 N.S.	3.0 1:116 N.S.	7.6 3:116 SIG.	2.1 1:109 N.S.	1.4 3:109 N.S.	1.2 3:109 N.S.	0 3:106 N.S.
SPACING	F D.F.	0 1:75 N.S.	11.5 1:75 SIG.	1.0 2:75 N.S.	0 1:75 N.S.	6.0 2:75 SIG.	11.3 2:75 SIG.	0 2:73 N.S.
ALTITUDE	F D.F.	0.1 1:65 N.S.	0.3 1:65 N.S.	5.5 2:65 SIG.	0.1 1:65 N.S.	0.5 2:65 N.S.	4.5 2:65 SIG.	0.7 2:63 N.S.
GROUND SPEED	F D.F.	0.1 1:65 N.S.	4.8 1:65 SIG.	0.7 2:65 N.S.	0.3 1:65 N.S.	0.3 2:65 N.S.	5.8 2:65 SIG.	0.9 2:63 N.S.
HEADING	F D.F.	6.5 1:61 SIG.	23.9 1:61 SIG.	32.4 2:61 SIG.	7.0 1:61 SIG.	3.2 2:61 SIG.	53.5 2:61 SIG.	0 2:59 N.S.

F ≡ F RATIO

D.F. ≡ DEGREES OF FREEDOM

N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)

SIG. ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

TABLE 3.3.3

CREW TREATMENT: RESULTS OF THREE WAY ANALYSIS OF VARIANCE FOR INFORMATION COMPONENTS

Information Component		MAIN EFFECTS			FIRST ORDER INTERACTIONS			SECOND ORDER INTERACTIONS
		CREW	SITUATION	AIRCRAFT SEQUENCE	CREW X SITUATION	CREW X AIRCRAFT SEQUENCE	SITUATION X AIRCRAFT SEQUENCE	CREW X SITUATION X AIRCRAFT SEQUENCE
POSITION	F D.F.	2.1 1:217 N.S.	1.7 3:217 N.S.	11.5 3:217 SIG.	0.2 3:202 N.S.	0.6 3:202 N.S.	0.8 9:202 N.S.	0.5 9:193 N.S.
SPACING	F D.F.	1.5 1:147 N.S.	19.6 3:147 SIG.	1.0 2:147 N.S.	≈0 3:147 N.S.	0.7 2:147 N.S.	10.2 6:147 SIG.	0.3 6:141 N.S.
ALTITUDE	F D.F.	0.1 1:133 N.S.	2 3:133 N.S.	0.2 2:133 N.S.	0.2 3:133 N.S.	1.2 2:133 N.S.	3.3 6:133 SIG.	0.5 6:127 N.S.
GROUND SPEED	F D.F.	0.4 1:144 N.S.	2.7 3:144 SIG	0.1 2:144 N.S.	0.6 3:144 N.S.	0.1 3:144 N.S.	1.5 6:144 N.S.	0.3 6:127 N.S.
HEADING	F D.F.	2.6 1:100 N.S.	22.6 3:100 SIG.	21.1 3:100 SIG.	1.5 3:100 N.S.	≈0 2:100 N.S.	32.7 6:100 SIG.	0.3 6:94 N.S.

F ≡ F RATIO

D.F. ≡ DEGREES OF FREEDOM

N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)

SIG. ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

TABLE 3.3.4

SPACING TASK TREATMENT: RESULTS OF FIRST LEVEL BREAKDOWN (BY SITUATION)
ANALYSIS OF VARIANCE FOR INFORMATION COMPONENTS

Information Components		SITUATION 2			SITUATION 3		
		Main Effects		First Order Interaction	Main Effects		First Order Interaction
		SPACING TASK	AIRCRAFT SEQUENCE	SPACING X AIRCRAFT SEQUENCE	SPACING TASK	AIRCRAFT SEQUENCE	SPACING X AIRCRAFT SEQUENCE
ALTITUDE	F D.F.	0.2 1:37 N.S.	5.6 2:37 SIG.	0.7 2:35 N.S.	0 1:30 N.S.	6.9 2:30 SIG.	0.1 2:28 N.S.
GROUND SPEED	F D.F.	0.6 1:37 N.S.	2.6 2:37 N.S.	0.6 2:35 N.S.	0 1:28 N.S.	4.4 2:23 SIG.	5.2 2:28 SIG.

F ≡ F RATIO

D.F. ≡ DEGREES OF FREEDOM

N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)

SIG ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

TABLE 3.3.5

CREW TREATMENT, RESULTS OF FIRST LEVEL BREAKDOWN (BY SITUATION) ANALYSIS OF VARIANCE FOR INFORMATION COMPONENTS

Information Components		SITUATION 1			SITUATION 2			SITUATION 3			SITUATION 4		
		Main Effects		First Order Inter.	Main Effects		First Order Inter.	Main Effects		First Order Inter.	Main Effects		First Order Inter.
		Crew	A/C	Crew X A/C	Crew	A/C	Crew X A/C	Crew	A/C	Crew X A/C	Crew	A/C	Crew X A/C
SPACING	F D.F.	1.8 1:38 N.S.	35.1 2:38 SIG	1.1 2:36 N.S.	0 1:33 N.S.	1.1 2:33 N.S.	3.1 2:31 N.S.	0.7 1:38 N.S.	4.1 2:38 SIG.	0.6 2:36 N.S.	0.1 1:40 N.S.	9.0 2:40 SIG.	2.1 2:38 N.S.
ALTITUDE	F D.F.	1.0 1:29 N.S.	1.0 2:29 N.S.	3.8 2:29 SIG	0 1:33 N.S.	2.8 2:33 N.S.	1.2 2:31 N.S.	1.6 1:34 N.S.	0.3 2:34 N.S.	1.8 2:32 N.S.	0.1 1:37 N.S.	2.7 2:37 N.S.	0 2:35 N.S.
HEADING	F D.F.	0.7 1:27 N.S.	0.3 2:27 N.S.	1.3 2:25 N.S.	0.1 1:21 N.S.	0.1 2:21 N.S.	≈0 2:19 N.S.	1.8 1:28 N.S.	285.1 2:28 SIG.	≈0 2:26 N.S.	2.3 1:26 N.S.	8.0 2:26 SIG.	0.3 2:24 N.S.

A/C ≡ AIRCRAFT

F ≡ F RATIO

D.F. ≡ DEGREES OF FREEDOM

N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)

SIG. ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

TABLE 3.3.6

ATC TREATMENT: RESULTS OF ANALYSIS OF VARIANCE FOR INFORMATION COMPONENTS

INFORMATION COMPONENTS		SITUATION 2				SITUATION 3			
		A/C +2	A/C +1	A/C -1	A/C 0	A/C +2	A/C +1	A/C -1	A/C 0
POSITION	F D.F.	1.8 1:19 N.S.	4.7 1:19 SIG.	0.3 1:13 N.S.	2.6 1:19 N.S.	0.2 1:15 N.S.	0 1:15 N.S.	40.6 1:17 SIG.	0.6 1:18 N.S.
SPACING	F D.F.	5.4 1:18 SIG.	12.1 1:18 SIG.	1.4 1:13 N.S.		0.4 1:14 N.S.			
ALTITUDE	F D.F.	0.9 1:15 N.S.	3.8 1:13 N.S.	0.4 1:14 N.S.		1.8 1:9 N.S.	0.3 1:14 N.S.	1.0 1:15 N.S.	
GROUND SPEED	F D.F.	3.7 1:14 N.S.	29.3 1:14 SIG.	0.2 1:14 N.S.		1.4 1:11 N.S.	4.7 1:13 SIG.	1.4 1:11 N.S.	
HEADING	F D.F.	2.9 1:15 N.S.	≈0 1:13 N.S.	0.1 1:9 N.S.		0.3 1:13 N.S.	2.4 1:14 N.S.	5.8 1:8 SIG.	

A/C ≡ AIRCRAFT F ≡ F RATIO D.F. ≡ DEGREES OF FREEDOM

N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$) SIG. ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

TABLE 3.3.7

SPACING TASK TREATMENT: RESULTS OF ANALYSIS OF VARIANCE FOR INFORMATION COMPONENTS

INFORMATION COMPONENTS		SITUATION 2				SITUATION 3			
		A/C +2	A/C +1	A/C -1	A/C 0	A/C +2	A/C +1	A/C -1	A/C 0
POSITION	F D.F.	0.5 1:116 N.S.							
SPACING	F D.F.	2.0 1:11 N.S.	1.1 1:14 N.S.	6.7 1:11 SIG.		0.2 1:12 N.S.	0.3 1:15 N.S.	≈0 1:10 N.S.	
ALTITUDE	F D.F.	0.1 1:37 N.S.				0 1:30 N.S.			
GROUND SPEED	F D.F.	0.6 1:37 N.S.				≈0 1:7 N.S.	0.5 1:14 N.S.	0.4 1:7 N.S.	
HEADING	F D.F.	0.9 1:11 N.S.	0.1 1:9 N.S.	0.8 1:7 N.S.		0.4 1:14 N.S.	5.2 1:13 SIG.	0.2 1:7 N.S.	

TABLE 3.3.8

CREW TREATMENT: RESULTS OF ANALYSIS OF VARIANCE FOR INFORMATION COMPONENTS

INFORMATION COMPONENTS		SITUATION 1				SITUATION 2				SITUATION 3				SITUATION 4			
		A/C +2	A/C +1	A/C -1	A/C 0	A/C +2	A/C +1	A/C -1	A/C 0	A/C +2	A/C +1	A/C -1	A/C 0	A/C +2	A/C +1	A/C -1	A/C 0
POSITION	F D.F.	2.1 1:217 N.S.															
SPACING	F D.F.	1.8 1:38 N.S.			X	0 1:33 N.S.			X	0.7 1:38 N.S.			X	0.1 1:40 N.S.			X
ALTITUDE	F D.F.	3.1 1:7 N.S.	0.4 1:12 N.S.	0.4 1:10 N.S.	X	0 1:33 N.S.			X	1.6 1:34 N.S.			X	0.1 1:37 N.S.			X
GROUND SPEED	F D.F.	0.4 * 1:144 N.S.															X
HEADING	F D.F.	0.7 1.27 N.S.			X	0.1 1:27 N.S.			X	1.8 1:28 N.S.			X	2.3 1:26 N.S.			X

F ≡ ? ratio

D.F. ≡ Degrees of Freedom

N.S. ≡ Not Statistically Significant $(p>.05)$

SlG. ≡ Statistically Significan: $(p<. 05)$

A/C ≡ AIRCRAFT

* ≡ A/C 0 NOT INCLUDED

of the hypotheses were refuted, then the contingency tables were broken down by the appropriate effect or effects. Tables 3.3.9 thru 3.3.11 present the results of the situation and aircraft hypothesis tests for the information point scores. Tables 3.3.12 and 3.3.13 present the results of Chi square tests for significance in point scores due to major treatment variable. In computing Chi Square, Yate's correction factor was used. Chi Square was judged to be significant if it reached the 5 percent level ($p<.05$).

The results of the situation and aircraft hypothesis tests for the null responses are presented in 3.3.14 thru 3.3.16. Table 3.3.17 shows the results of the Chi Square tests performed on the null responses for the major treatment variable effects. Again Yate's correction factor was used in computing Chi Square and significance was judged at the five percent level ($p<.05$).

3.3.2 Conflict Detection

Pilot reaction times to ILS crossover intrusions were measured for the crew treatment variable only. All non-alarm single pilots were analyzed for the effect of information update rate on conflict reaction time in a two way analysis of variance. The main treatments in this analysis were update rate and situation. The results of this analysis are presented in Table 3.3.18.

TABLE 3.3.9

ATC TREATMENT: RESULTS OF SITUATION AND AIRCRAFT HYPOTHESES TESTS FOR LANDING SEQUENCE INFORMATION COMPONENT

(N.S.) : NOT QUITE STATISTICALLY SIGNIFICANT

ATC TREATMENT	FUTURE SYSTEM	PRESENT SYSTEM
χ^2	0.47	3.52
DEGREES OF FREEDOM	1	1
p	.5 N.S.	.07 (N.S.)

TEST OF SITUATION HYPOTHESIS (SITUATIONS 2 AND 3)

(a)

ATC TREATMENT	FUTURE SYSTEM	PRESENT SYSTEM
χ^2	0.15	1.60
DEGREES OF FREEDOM	2	[illegible]
p	.92 N.S.	.2 N.S.

TEST OF AIRCRAFT HYPOTHESIS (AIRCRAFTS +2, +1, AND -1)

(b)

$\chi^2 \equiv$ CHI SQUARE

p $\equiv$ PROBABILITY χ^2 COULD BE EXCEEDED BY CHANCE

TABLE 3.3.10

SPACING TASK TREATMENT: RESULTS OF SITUATION AND AIRCRAFT HYPOTHESIS TESTS FOR LANDING SEQUENCE INFORMATION COMPONENT

SPACING TASK TREATMENT	NO SPACING TASK	SPACING TASK
χ^2	0.47	.15
DEGREES OF FREEDOM	1	1
p	.5 N.S.	.7 N.S.

TEST OF SITUATION HYPOTHESIS (SITUATIONS 2 AND 3)

(c)

SPACING TASK TREATMENT	NO SPACING TASK	SPACING TASK
χ^2	0.15	.94
DEGREES OF FREEDOM	2	1
p	.92 N.S.	.32 N.S.

TEST OF AIRCRAFT HYPOTHESIS (AIRCRAFTS +2, +1, AND -1)

(d)

TABLE 3.3.11

CREW TREATMENT: RESULTS OF SITUATION AND AIRCRAFT HYPOTHESES TESTS

CREW TREATMENT	SINGLE PILOTS		CREW	
INFORMATION COMPONENT	IDENTI- FICATION	LANDING SEQUENCE	IDENTI- FICATION	LANDING SEQUENCE
χ^2	0.99	0.69	15.30	0.05
DEGREES OF FREEDOM	3	3	3	3
P	.82 N.S.	.89 N.S.	.01 SIG.	.99 N.S.

TEST OF SITUATION HYPOTHESIS (SITUATIONS 1, 2, 3, AND 4)

(a)

CREW TREATMENT	SINGLE PILOTS		CREW	
INFORMATION COMPONENTS	IDENTI- FICATION	LANDING SEQUENCE	IDENTI- FICATION	LANDING SEQUENCE
χ^2	0.58	0.37	0.27	0.33
DEGREES OF FREEDOM	2	2	2	2
P	.75 N.S.	.85 N.S.	.9 N.S.	.85 N.S.

TEST OF AIRCRAFT HYPOTHESIS (AIRCRAFT +2, +1, and -1)

(b)

$\chi^2 \equiv$ CHI SQUARE

p $\equiv$ PROBABILITY χ^2 COULD BE EXCEEDED BY CHANCE

N.S. $\equiv$ NOT STATISTICALLY SIGNIFICANT $(p > .05)$

SIG. $\equiv$ STATISTICALLY SIGNIFICANT $(p < .05)$

TABLE 3.3.12

RESULTS OF CREW TREATMENT TEST FOR IDENTIFICATION INFORMATION COMPONENT

		SITUATION 1			SITUATION 2			SITUATION 3			SITUATION 4		
		AIRCRAFT +2	AIRCRAFT +1	AIRCRAFT -1	AIRCRAFT +2	AIRCRAFT +1	AIRCRAFT -1	AIRCRAFT +2	AIRCRAFT +1	AIRCRAFT -1	AIRCRAFT +2	AIRCRAFT +1	AIRCRAFT -1
CREW TREATMENT	χ^2	0.02	0.07	0.09	0.59	0.23	2.1	0.23	0.03	0.42	8.41	2.11	0.03
	D.F.	1	1	1	1	1	1	1	1	1	1	1	1
	p	.88	.3	.3	.48	.6	.15	.6	.85	.5	.01	.15	.85
		N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	SIG.	N.S.	N.S.

TABLE 3.3.13

RESULTS OF CREW, ATC, AND SPACING TASK TREATMENT TEST, LANDING SEQUENCE INFORMATION COMPONENT

TREATMENT CONDITION		LANDING SEQUENCE
CREW TREATMENT	χ^2	2.88
	D.F.	1
	p	.1
		N.S.
ATC TREATMENT	χ^2	16.96
	D.F.	1
	p	.01
		SIG.
SPACING TASK TREATMENT	χ^2	2.12
	D.F.	1
	p	.15
		N.S.

χ^2 = CHI SQUARE
D.F. = DEGREES OF FREEDOM
N.S. = NOT STATISTICALLY SIGNIFICANT ($p > .05$)
SIG. = STATISTICALLY SIGNIFICANT ($p < .05$)

TABLE 3.3.14

ATC TREATMENT: SITUATION AND AIRCRAFT HYPOTHESIS TEST FOR NULL RESPONSES

ATC TREATMENT	FUTURE SYSTEM				TODAY'S SYSTEM			
INFORMATION COMPONENTS	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING
χ^2	0.20	0.82	0.31	0.18	4.79	0.37	0.24	1.60
DEGREES OF FREEDOM	1	1	1	1	1	1	1	1
P	.65 N.S.	.4 N.S.	.72 N.S.	.65 N.S.	.03 SIG.	.55 N.S.	.65 N.S.	.2 N.S.

TEST OF SITUATION HYPOTHESIS (SITUATIONS 2 AND 3)

(a)

ATC TREATMENT	FUTURE SYSTEM				TODAY'S SYSTEM			
INFORMATION COMPONENTS	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING
χ^2	1.90	0.73	2.00	0.09	3.51	0.61	0.07	1.80
DEGREES OF FREEDOM	2	2	2	2	2	2	2	2
P	.4 N.S.	.7 N.S.	.35 N.S.	.95 N.S.	.15 N.S.	.75 N.S.	.95 N.S.	.4 N.S.

TEST OF AIRCRAFT HYPOTHESIS (AIRCRAFT +2, +1, AND -1)

(b)

χ^2 ≡ CHI SQUARE
p ≡ PROBABILITY THAT χ^2 COULD BE EXCEEDED BY CHANCE
N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)
SIG ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

TABLE 3.3.15

SPACING TASK TREATMENT: SITUATION AND AIRCRAFT HYPOTHESIS TEST FOR NULL RESPONSES

SPACING TREATMENT	NO SPACING TASK				SPACING TASK			
INFORMATION COMPONENT	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING
χ^2	0.20	0.82	0.11	0.18	1.43	1.74	1.95	0.67
DEGREES OF FREEDOM	1	1	1	1	1	1	1	1
p	.65 N.S.	.4 N.S.	.72 N.S.	.65 N.S.	.25 N.S.	.18 N.S.	.15 N.S.	.4 N.S.

TEST OF SITUATION HYPOTHESIS (SITUATIONS 2 AND 3)

(a)

SPACING TREATMENT	NO SPACING TASK				SPACING TASK			
INFORMATION COMPONENTS	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING
χ^2	1.90	0.73	2.00	0.09	5.01	8.00	5.82	6.14
DEGREES OF FREEDOM	2	2	2	2	2	2	2	2
p	.4 N.S.	.7 N.S.	.35 N.S.	.95 N.S.	.08 (N.S.)	.02 SIG.	.06 (N.S.)	.05 SIG.

TEST OF AIRCRAFT HYPOTHESIS (AIRCRAFT +2, +1, AND -1)

(b)

χ^2 ≡ CHI SQUARE

p ≡ PROBABILITY THAT χ^2 COULD BE EXCEEDED BY CHANCE

N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)

(N.S.) ≡ NOT QUITE STATISTICALLY SIGNIFICANT

SIG. ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

TABLE 3.3.16

CREW TREATMENT. SITUATION AND AIRCRAFT HYPOTHESIS TEST FOR NULL RESPONSES

CREW TREATMENT	SINGLE PILOTS				CREW			
INFORMATION COMPONENT	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING
χ^2	3.00	3.60	3.70	0.60	3.00	2.00	2.00	0.50
Degrees of Freedom	3	3	3	3	3	3	3	3
p	.4 N.S.	.4 N.S.	.3 N.S.	.9 N.S.	.4 N.S.	.6 N.S.	.6 N.S.	.9 N.S.

TEST OF SITUATION HYPOTHESIS (SITUATIONS 1, 2, 3, AND 4)

(a)

CREW TREATMENT	SINGLE PILOTS				CREW			
INFORMANT COMPONENT	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING	POSITION-SPACING	ALTITUDE	GROUND SPEED	HEADING
χ^2	1.73	3.00	3.25	0.14	2.02	1.00	1.00	0.88
Degrees of Freedom	2	2	2	2	2	2	2	2
p	.4 N.S.	.4 N.S.	.35 N.S.	.92 N.S.	.35 N.S.	.8 N.S.	.8 N.S.	.65 N.S.

TEST OF AIRCRAFT HYPOTHESIS (AIRCRAFT +2, +1, AND -1)

(b)

χ^2 ≡ CHI SQUARE
p ≡ PROBABILITY THAT χ^2 COULD BE EXCEEDED BY CHANCE
N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)
SIG. ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

TABLE 3.3.17

RESULTS OF CREW, ATC, AND SPACING TASK TREATMENT HYPOTHESES TESTS FOR NULL RESPONSES

		POSITION-SPACING			ALTITUDE			GROUND SPEED			HEADING		
		AIRCRAFT +2	AIRCRAFT +1	AIRCRAFT -1	AIRCRAFT +2	AIRCRAFT +1	AIRCRAFT -1	AIRCRAFT +2	AIRCRAFT +1	AIRCRAFT -1	AIRCRAFT +2	AIRCRAFT +1	AIRCRAFT -1
CREW CONDITION TREATMENT	χ^2 D.F. p	1.05 1 .3 N.S.			7.12 1 .01 SIG			7.12 1 .01 SIG			0.73 1 .4 N.S.		
ATC TREATMENT	χ^2 D.F. p	Sit. 2:.83 Sit.3:.85 Sit. 2: 1 Sit.3: 1 Sit. 2:.4 Sit.3:.4 Sit. 2:N.S.Sit.3:N.S.			3.55 1 .06 (N.S.)			10.11 1 .01 SIG			2.96 1 .09 (N.S.)		
SPACING TASK TREATMENT	χ^2 D.F. p	3.05 1 .08 (N.S.)	0.02 1 .68 N.S.	0.76 1 .4 N.S.	0.95 1 .3 N.S.	0.05 1 .8 N.S.	6.98 1 .01 SIG.	5.95 1 .02 SIG.	0.07 1 .8 N.S.	5.95 1 .02 SIG.	0.35 1 .55 N.S.	1.12 1 .3 N.S.	0.69 1 .4 N.S.

TEST FOR TREATMENT CONDITION HYPOTHESES

χ^2 ≡ CHI SQUARE
D.F. ≡ DEGREES OF FREEDOM
p ≡ PROBABILITY THAT χ^2 COULD BE EXCEEDED BY CHANCE
N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)
(N.S.) ≡ NOT QUITE STATISTICALLY SIGNIFICANT
SIG. ≡ STATISTICALLY SIGNIFICANT ($p < .05$)
Sit. ≡ SITUATION

TABLE 3.3.18

UPDATE RATE TREATMENT: RESULTS OF TWO WAY ANALYSIS OF VARIANCE FOR CONFLICT DETECTION REACTION TIMES (SINGLE PILOTS-NO ALARM ONLY)

	MAIN EFFECTS		FIRST ORDER INTERACTION
	UPDATE RATE	SITUATION	UPDATE RATE X SITUATION
F	0.3	18.7	0.4
D.F.	1:13	1:13	1:12
	N.S.	SIG.	N.S.

F ≡ F RATIO
D.F. ≡ DEGREES OF FREEDOM
N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)
SIG. ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

A three way analysis of variance was performed on the reaction time data to test for crew, alarm, and situation effects. The same procedure concerning breakdown analysis used in the information component analyses was used here. Since the data showed no statistically significant update rate effect, both one second and four second information update rate data have been pooled. Tables 3.3.19 and 3.3.20 present the results of this analysis. The F ratio test was judged to be significant if the probability of a chance result was below the five percent level ($p<.05$).

Also of interest is whether or not a conflict was detected during single runway and parallel runway simulations under all applicable major treatment conditions. As previously stated, no parallel runway tests were conducted comparing the air traffic control display/communication treatments. In addition, the single runway tests for both air traffic control and spacing task treatments involved only one conflict scenario. For this reason, the situation null hypothesis was only tested for the parallel runway simulations of the spacing task treatment and both the single and parallel runway scenarios of the crew treatment. The results of these situation hypothesis tests are shown in Tables 3.3.21 and 3.3.22.

The results of all three of the major treatment variables is presented in Table 3.3.23. The results are broken down by

TABLE 3.3.19

CREW AND ALARM TREATMENT: RESULTS OF THREE WAY ANALYSIS OF VARIANCE FOR CONFLICT DETECTION REACTION TIMES

	MAIN EFFECTS			FIRST ORDER INTERACTIONS			SECOND ORDER INTERACTIONS
	CREW	ALARM	SITUATION	CREW X ALARM	CREW X SITUATION	ALARM X SITUATION	CREW X ALARM X SITUATION
F	2.4	11.0	16.4	0.6	1.1	8.7	0
D.F.	1:23	1:23	1:23	1:23	1:23	1:23	1:22
	N.S.	SIG.	SIG.	N.S.	N.S.	SIG.	N.S.

TABLE 3.3.20

CREW AND ALARM TREATMENT: RESULTS OF BREAKDOWN BY SITUATION ANALYSIS OF VARIATION FOR CONFLICT DETECTION REACTION TIMES

	SITUATION 6			SITUATION 7		
	MAIN EFFECTS		FIRST ORDER INTERACTION	MAIN EFFECTS		FIRST ORDER INTERACTION
	CREW	ALARM	CREW X ALARM	CREW	ALARM	CREW X ALARM
F	0.7	16.6	2.7	2.9	0.2	0.1
D.F.	1:10	1:10	1:9	1:14	1.14	1:13
	N.S.	SIG.	N.S.	N.S.	N.S.	N.S.

F ≡ F RATIO
D.F. ≡ DEGREES OF FREEDOM
N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)
SIG. ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

TABLE 3.3.21

SPACING TASK TREATMENT: SITUATION HYPOTHESIS TEST (SITUATIONS 6 and 7, WITH SPACING TASK SUBJECTS ONLY)

RUNWAY CONFIGURATION	PARALLEL RUNWAY
DETECTION POINT REFERENCE	AT OR BEFORE CPA
χ^2 D.F. P	.82 1 .4 N.S.*

TABLE 3.3.22

CREW TREATMENT: SITUATION HYPOTHESIS TEST (SINGLE RUNWAY - SITUATION 3 and 4, PARALLEL RUNWAY - SITUATION 6 and 7)

RUNWAY CONFIGURATION	SINGLE RUNWAY							PARALLEL RUNWAY				
DETECTION POINT REFERENCE	AT OR BEFORE STOP-ACTION QUIZ		AT OR BEFORE CLOSEST POINT OF APPROACH					AT OR BEFORE CLOSEST POINT OF APPROACH				
CREW TREATMENT	SINGLE PILOT	CREW	SINGLE PILOT		CREW			SINGLE PILOT			CREW	
ALARM TREATMENT	A & $\bar{A}$	A & $\bar{A}$	A	$\bar{A}$	A	$\bar{A}$		A	$\bar{A}$		A	$\bar{A}$
UPDATE RATE TREATMENT	4 SECOND	4 SECOND	4 SEC.	4 SEC.	4 SEC.	4 SEC.		4 SEC.	1 SEC.	4 SEC.	4 SEC.	4 SEC.
χ^2 D.F. P	0 1 1.0 N.S.*	0.33 1 .6 N.S.	0 1 1.0 N.S.	0 1 1.0 N.S.	0 1 1.0 N.S.	0 1 1.0 N.S.		0 1 1.0 N.S.	0 1 1.0 N.S.	0 1 1.0 N.S.	0 1 1.0 N.S.	0 1 1.0 N.S.

A ≡ ALARM, $\bar{A}$ ≡ NO ALARM, χ^2 ≡ CHI SQUARE, D.F. ≡ DEGREES OF FREEDOM

*N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p>.05$), SIG.≡ STATISTICALLY SIGNIFICANT ($p<.05$)

TABLE 3.3.23

TEST FOR TREATMENT CONDITION HYPOTHESIS

MAIN TREATMENT	CREW TREATMENT					SPACING TASK TREATMENT			ATC TREATMENT	
RUNWAY CONFIGUR-ATION	SINGLE RUNWAY			PARALLEL RUNWAY		SINGLE RUNWAY		PARALLEL RUNWAY	SINGLE RUNWAY	
DETECTION POINT REFERENCE	AT OR BEFORE SAQ	AT OR BEFORE CPA		AT OR BEFORE CPA		AT OR BEFORE SAQ	AT OR BEFORE CPA	AT OR BEFORE CPA	AT OR BEFORE SAQ	AT OR BEFORE CPA
ALARM TREATMENT	A & $\bar{A}$	A	$\bar{A}$	A	$\bar{A}$	$\bar{A}$	$\bar{A}$	$\bar{A}$	$\bar{A}$	$\bar{A}$
χ^2	0	0	0	0	0	0.59	1.07	1.8	2.17	8.96
D.F.	1	1	1	1	1	1	1	1	1	1
p	1.0	1.0	1.0	1.0	1.0	.4	.3	.18	.15	.01
	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	SIG.

TABLE 3.3.24

EMPIRICAL PROBABILITY OF DETECTION AT OR BEFORE THE CLOSEST POINT OF APPROACH

MAJOR TREATMENT VARIABLE	SINGLE PILOT (FUTURE SYSTEM, NO SPACING TASK)				CREW				SPACING TASK		TODAY'S SYSTEM
RUNWAY CONFIGUR-ATION	SINGLE RUNWAY		PARALLEL RUNWAY		SINGLE RUNWAY		PARALLEL RUNWAY		SINGLE RUNWAY	PARALLEL RUNWAY	SINGLE RUNWAY
ALARM TREAT-MENT	A	$\bar{A}$	A	$\bar{A}$	A	$\bar{A}$	A	$\bar{A}$	$\bar{A}$	$\bar{A}$	$\bar{A}$
$\hat{P}$ (D)	1	1	1	1	1	1	1	1	.63	.79	.1
NO. OF SAMPLES	(8)	(12)	(8)	(16)	(6)	(6)	(5)	(5)	(8)	(14)	(10)

A ≡ ALARM $\bar{A}$ ≡ NO ALARM χ^2 ≡ CHI SQUARE   SAQ ≡ STOP-ACTION QUIZ
CPA ≡ CLOSEST POINT OF APPROACH   D.F. ≡ DEGREES OF FREEDOM
p ≡ PROBABILITY χ^2 COULD BE EXCEEDED BY CHANCE $\hat{P}$(D) ≡ EMPIRICAL PROBABILITY OF DETECTION AT OR BEFORE CPA   N.S. ≡ NOT STATISTICALLY SIGNIFICANT ($p > .05$)
SIG. ≡ STATISTICALLY SIGNIFICANT ($p < .05$)

runway configuration (single or parallel), detection point reference time (at or before stop-action quiz or at or before closest point of approach), and alarm treatment (no alarm or alarm). A Chi Square satisfying the five percent ($p<.05$) significance level was considered statistically significant.

Empirical probabilities of detection based on the number of observed detections divided by the number of possible detections were computed for each of the treatments considered above. These detection probabilities are presented in Table 3.3.24.

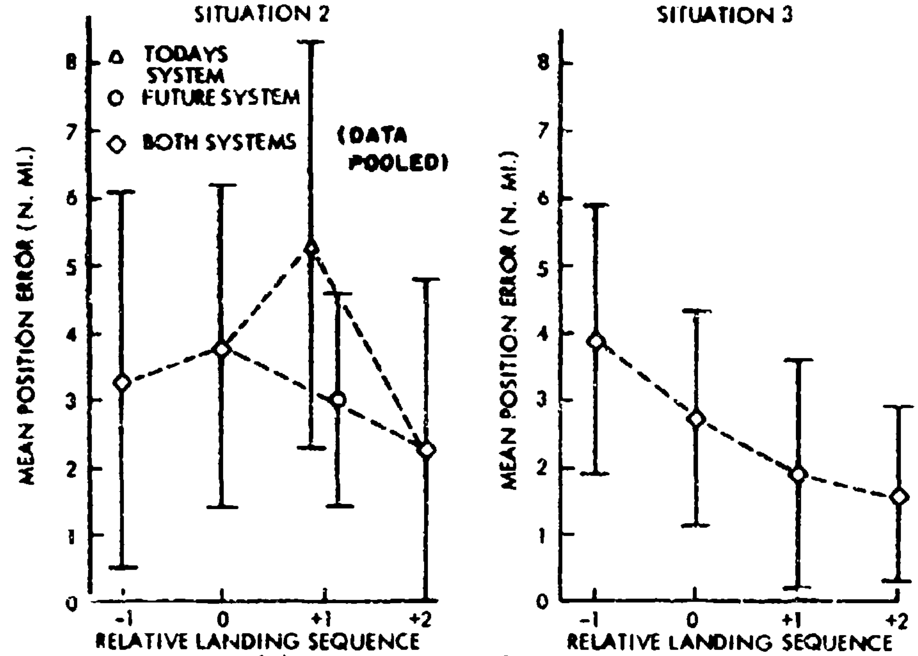
3.4 Post-Analysis Data Summary

3.4.1 Information Components

Results of the analysis of variance for the information components were used to guide the combination of data from the main effect conditions. If no statistical significance was found for the major treatment variable, situation, or aircraft sequence treatment, then the data for the non-significant treatment were pooled.

Combined graphs of the information component errors are presented in Figures 3.4.1 thru 3.4.15. These graphs depict the means values plus and minus one standard deviation. The signed mean rather than the absolute value of the mean is shown. Where a major treatment variable showed significance for given aircraft, separate component estimates are indicated.

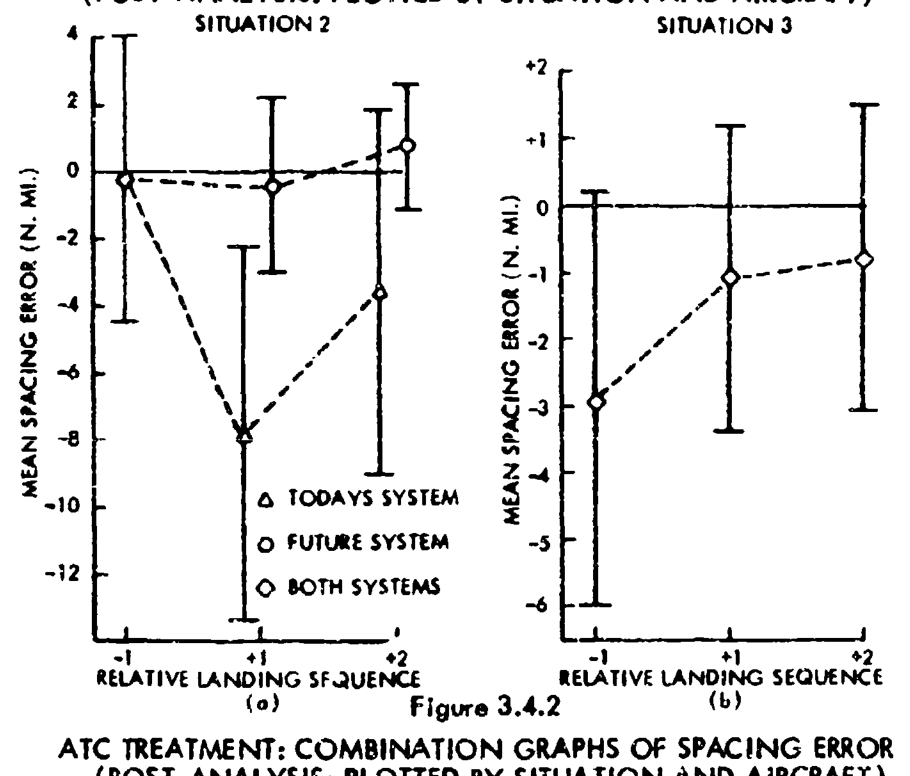
Histograms of null responses for each major treatment variable are presented in Figures 3.4.16 thru 3.4.18. The



(a) (b)

Figure 3.4.1

ATC TREATMENT: COMBINATION GRAPHS OF POSITION ERROR (POST-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)



(a) (b)

Figure 3.4.2

ATC TREATMENT: COMBINATION GRAPHS OF SPACING ERROR (POST-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

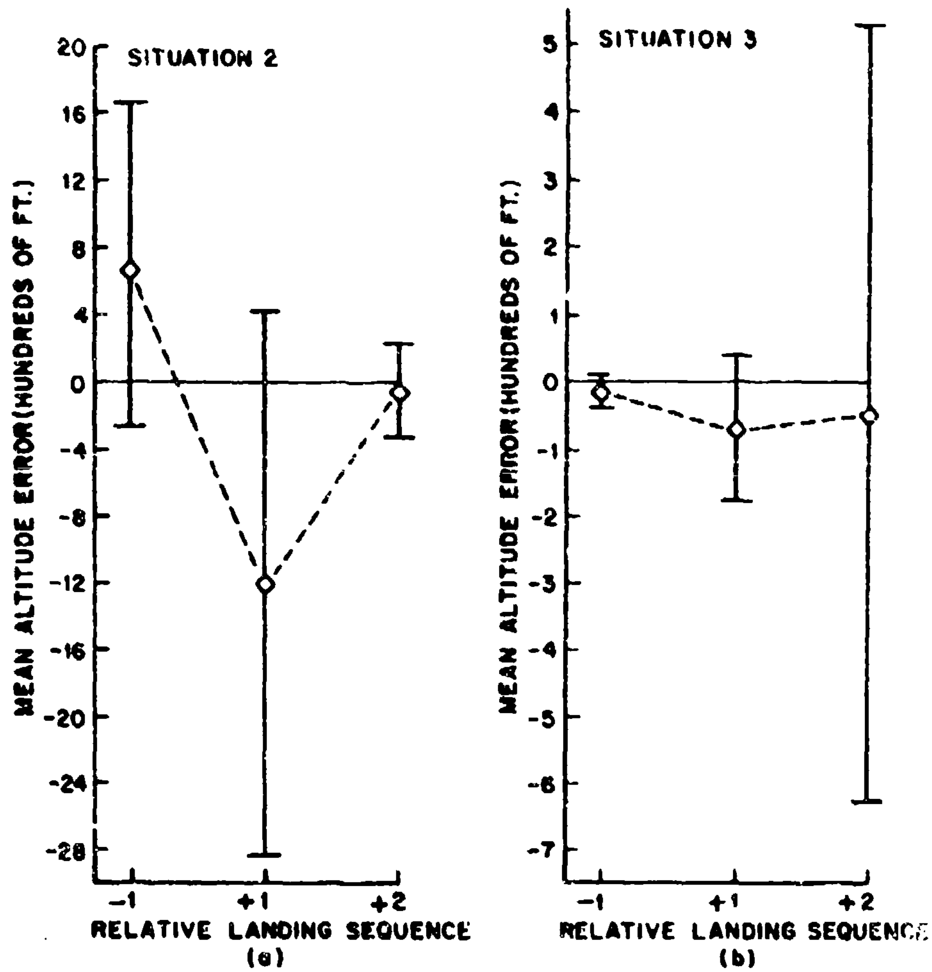


Figure 3.4.3

ATC TREATMENT: COMBINATION GRAPHS OF ALTITUDE ERROR
(POST-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

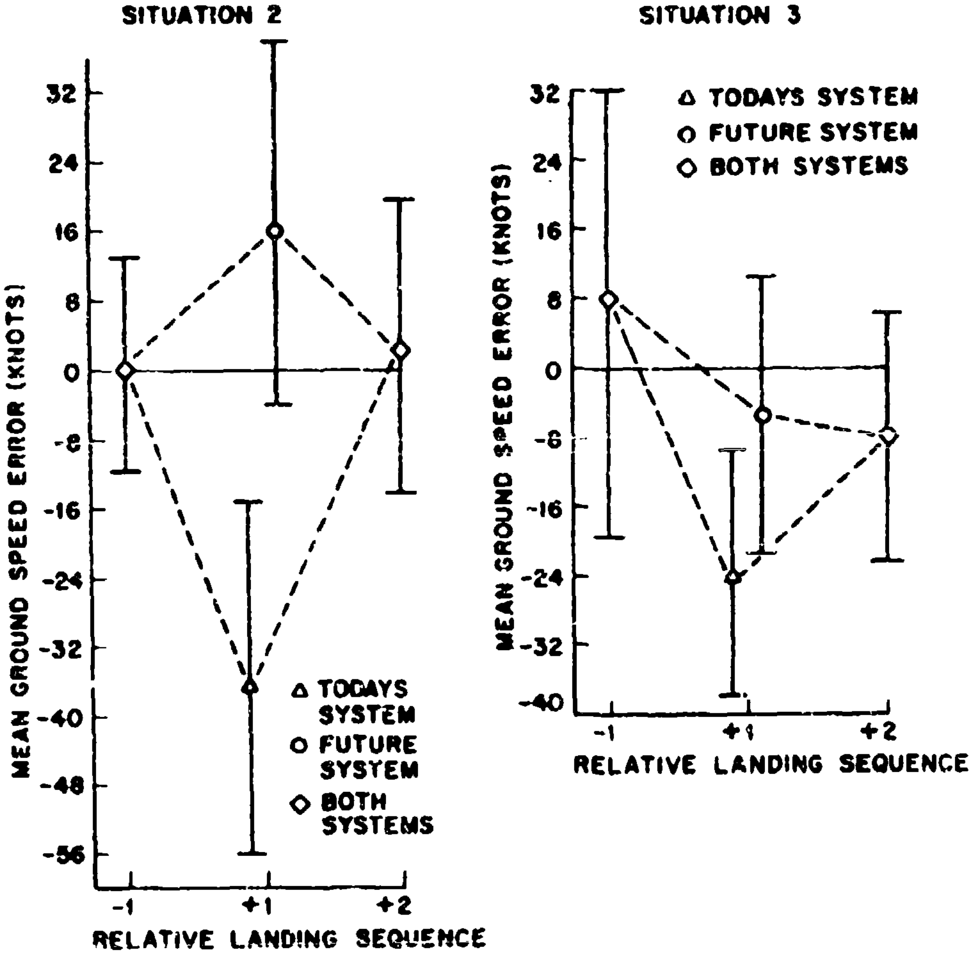


Figure 3.4.4

ATC TREATMENT: COMBINATION GRAPHS OF GROUND SPEED ERROR
(POST-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

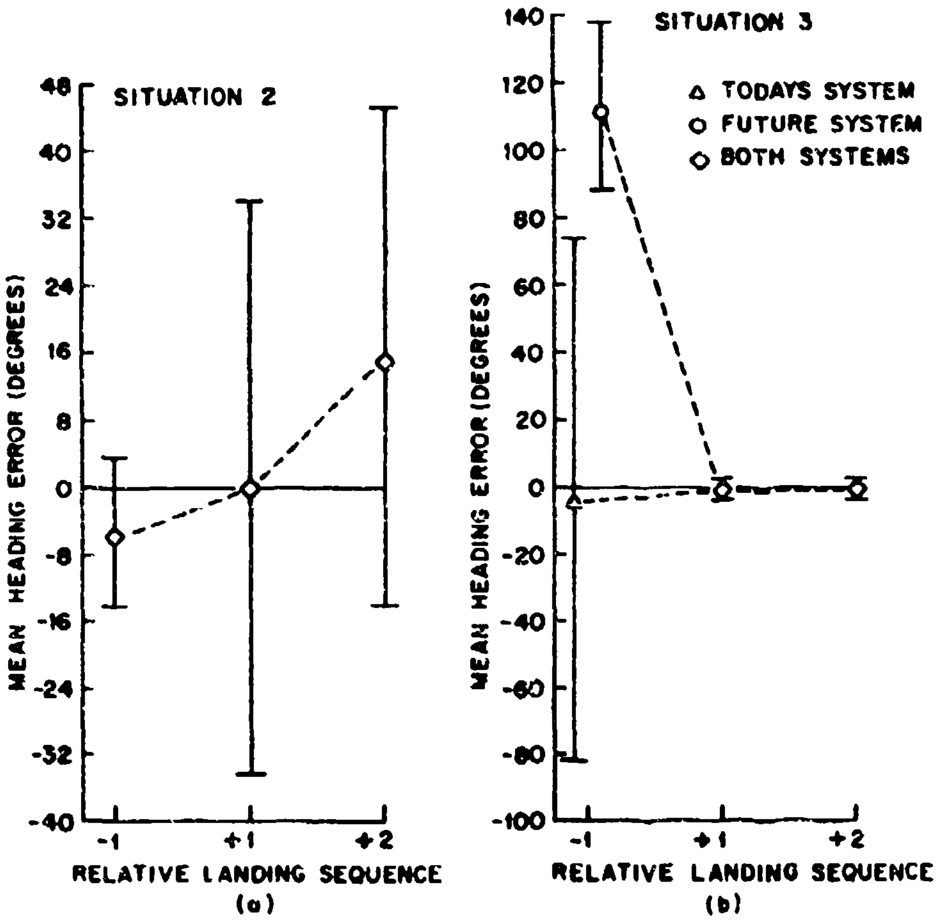


Figure 3.4.5

ATC TREATMENT: COMBINATION GRAPHS OF HEADING ERROR
(POST-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

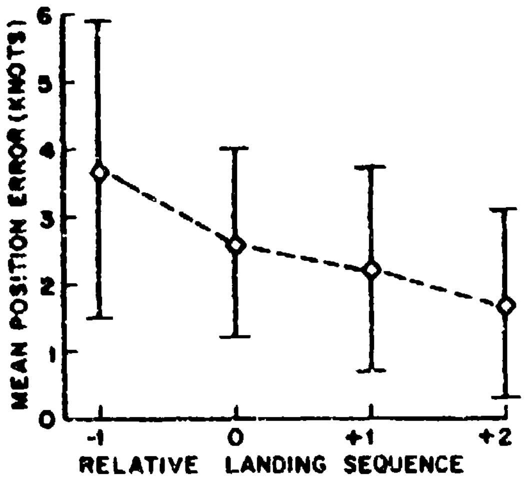


Figure 3.4.6

SPACING TASK TREATMENT: COMBINATION GRAPH OF POSITION ERROR
(POST-ANALYSIS: PLOTTED BY AIRCRAFT)

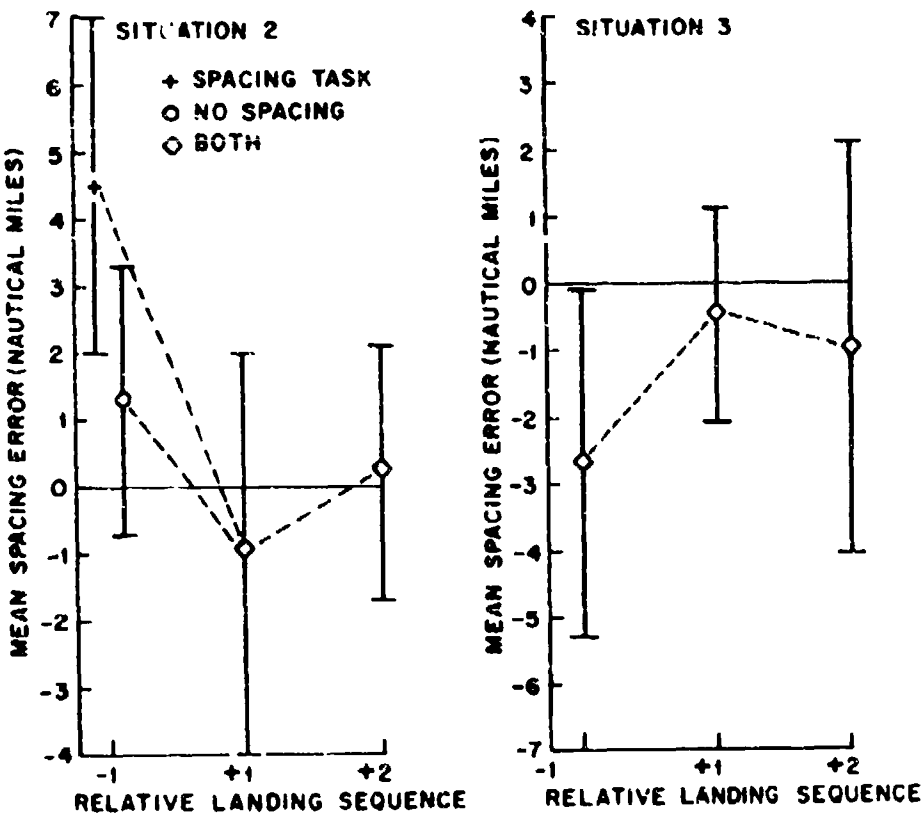


Figure 3.4.7

SPACING TASK TREATMENT: COMBINATION GRAPHS OF SPACING ERROR
(POST-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

SPACING TASK TREATMENT: COMBINATION GRAPHS OF ALTITUDE ERROR
(POST-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

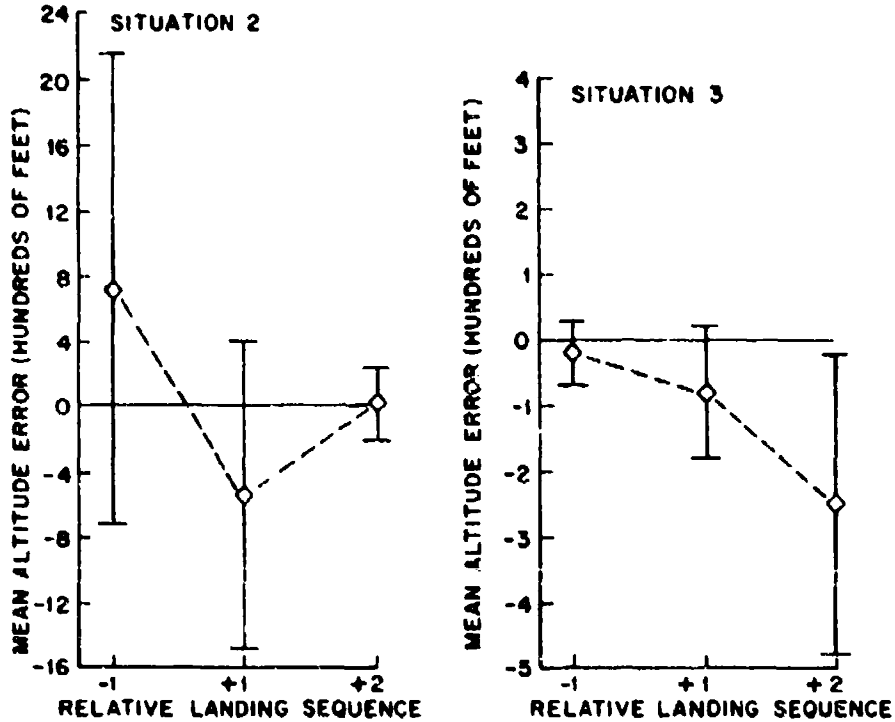


Figure 3.4.8

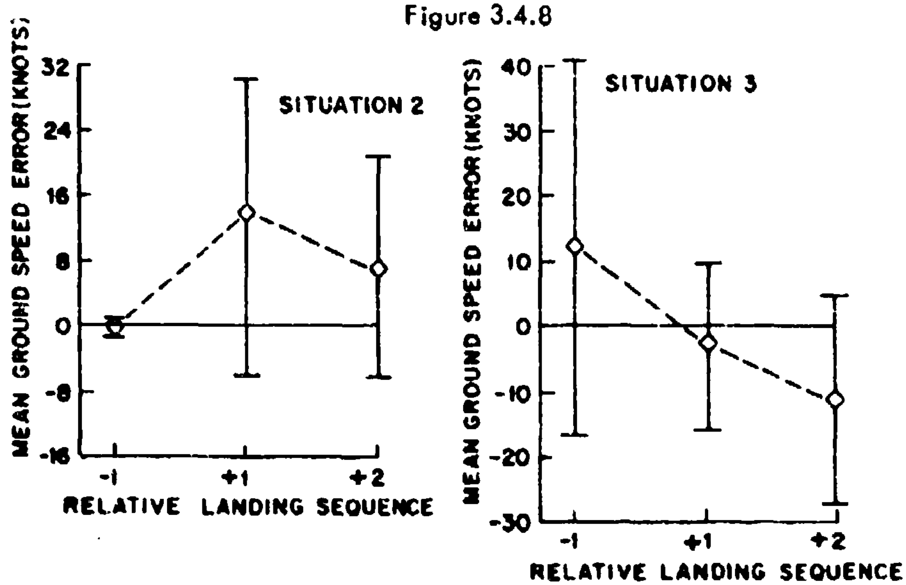


Figure 3.4.9

SPACING TASK TREATMENT: COMBINATION GRAPHS OF GROUND SPEED ERROR
(POST-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

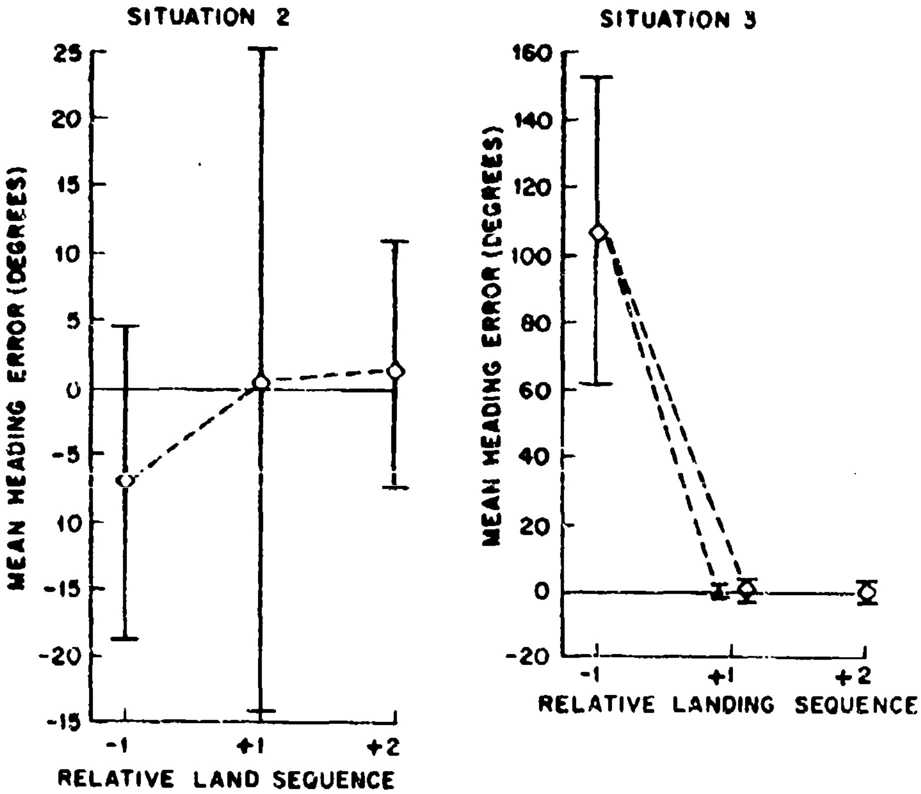


Figure 3.4.10

SPACING TASK TREATMENT: COMBINATION GRAPHS OF HEADING ERROR (POST-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

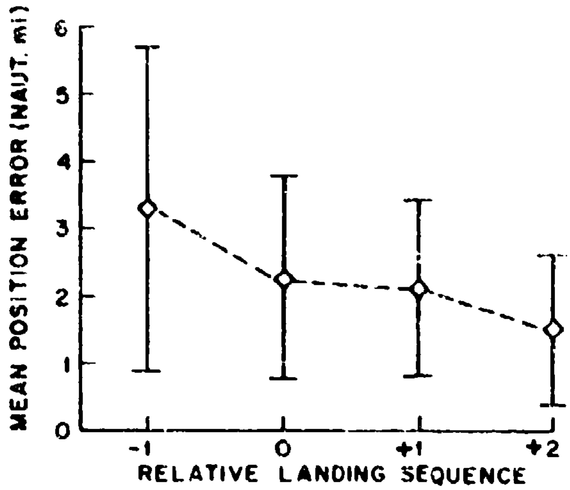


Figure 3.4.11

CREW TREATMENT: COMBINATION GRAPH OF POSITION ERROR
(POST-ANALYSIS: PLOTTED BY AIRCRAFT)

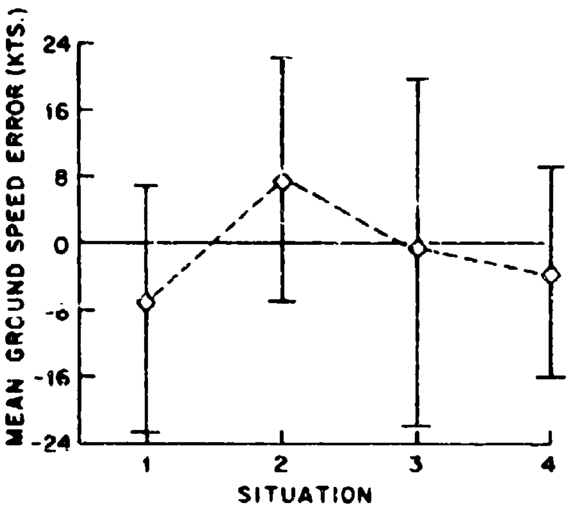


Figure 3.4.12

CREW TREAMENT: COMBINATION GRAPH OF GROUND SPEED ERROR
(POST-ANALYSIS: PLOTTED BY SITUATION)

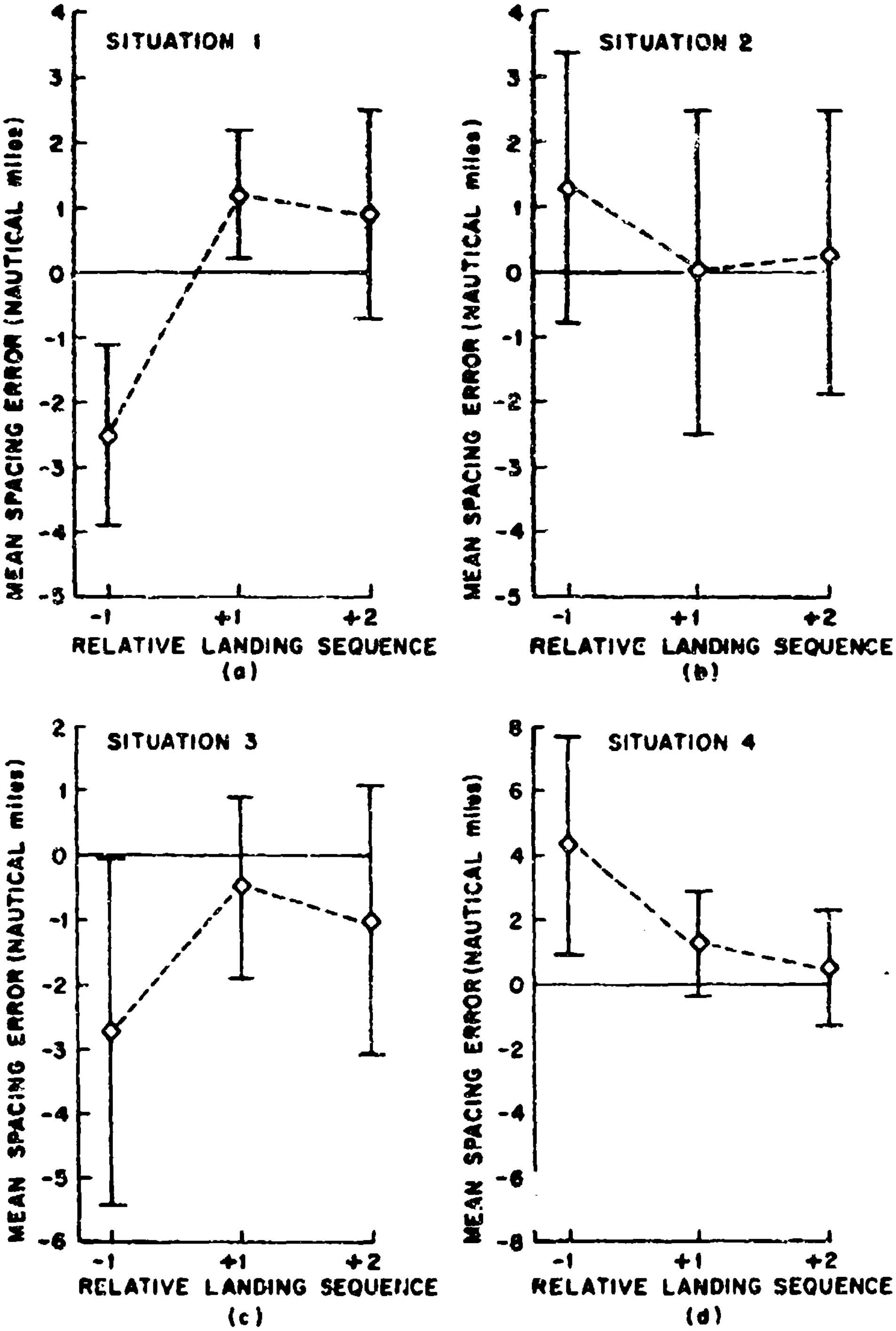


Figure 3.4.13

CREW TREATMENT: COMBINATION GRAPHS OF SPACING ERROR
(POST ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

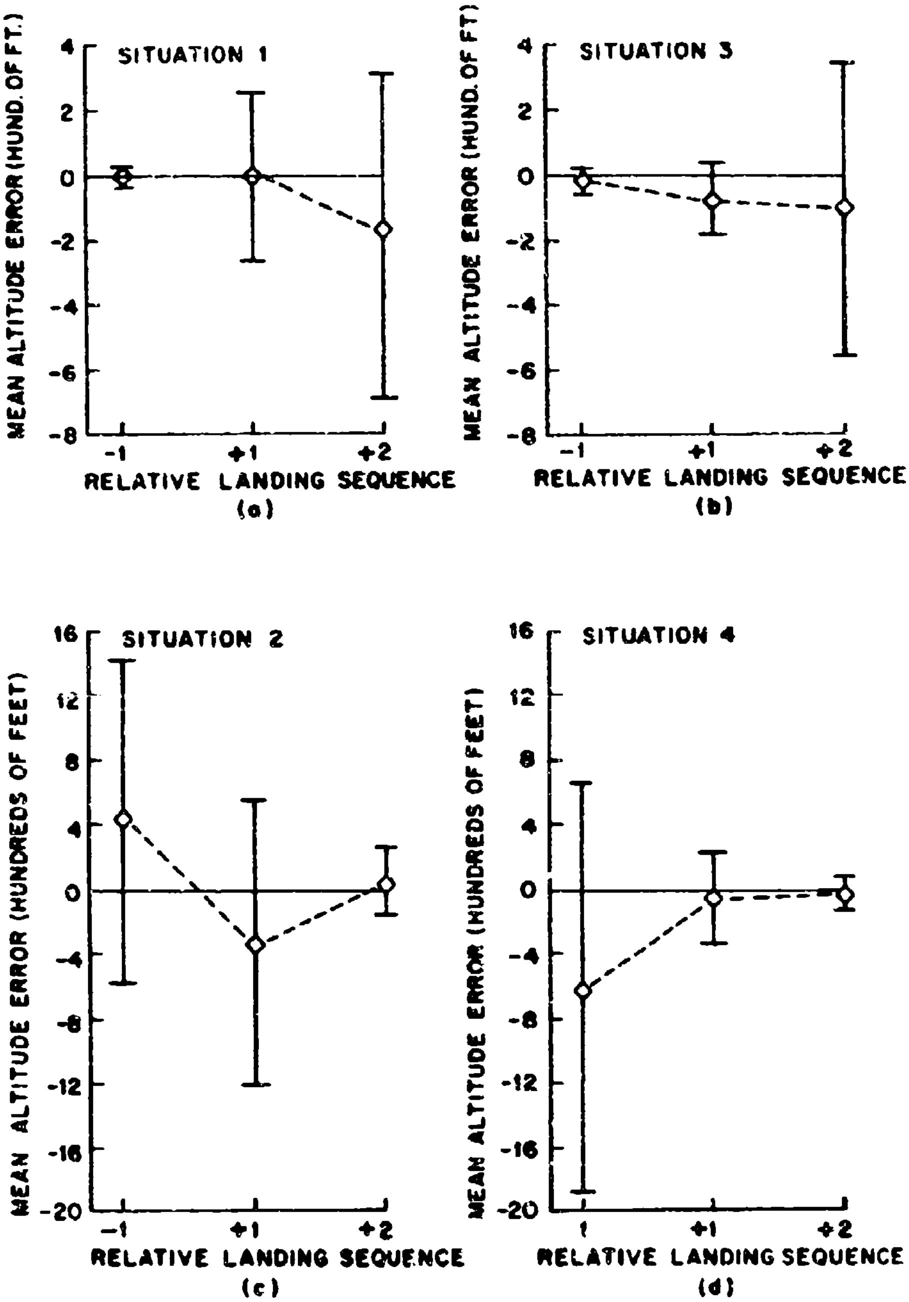


Figure 3.4.14

CREW TREATMENT: COMBINATION GRAPHS OF ALTITUDE ERROR (POST-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

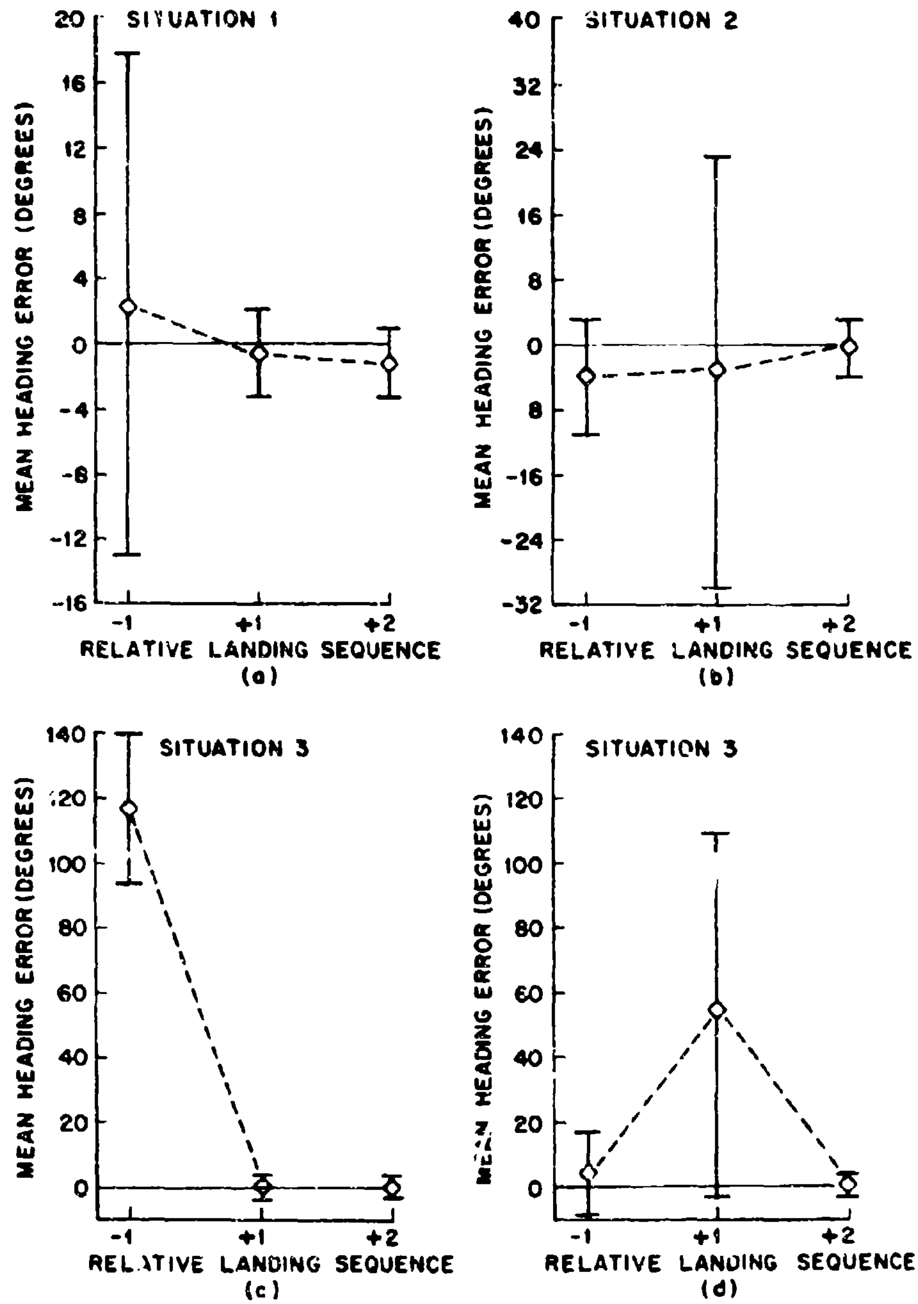


Figure 3.4.15

CREW TREATMENT: COMBINATION GRAPHS OF HEADING ERROR (POST-ANALYSIS: PLOTTED BY SITUATION AND AIRCRAFT)

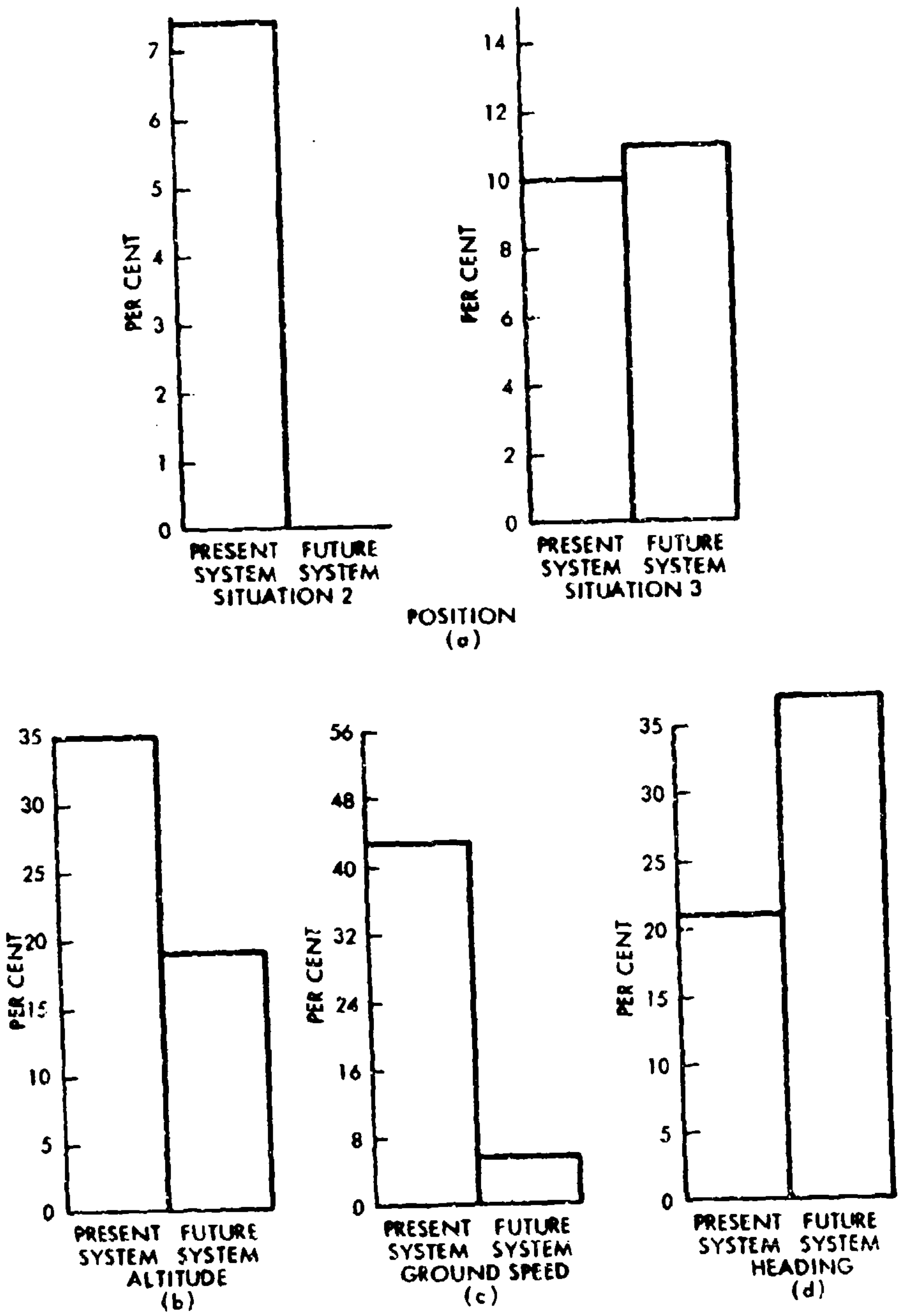


Figure 3.4.16

ATC TREATMENT: NULL RESPONSES OF INFORMATION COMPONENTS (PRE-ANALYSIS)

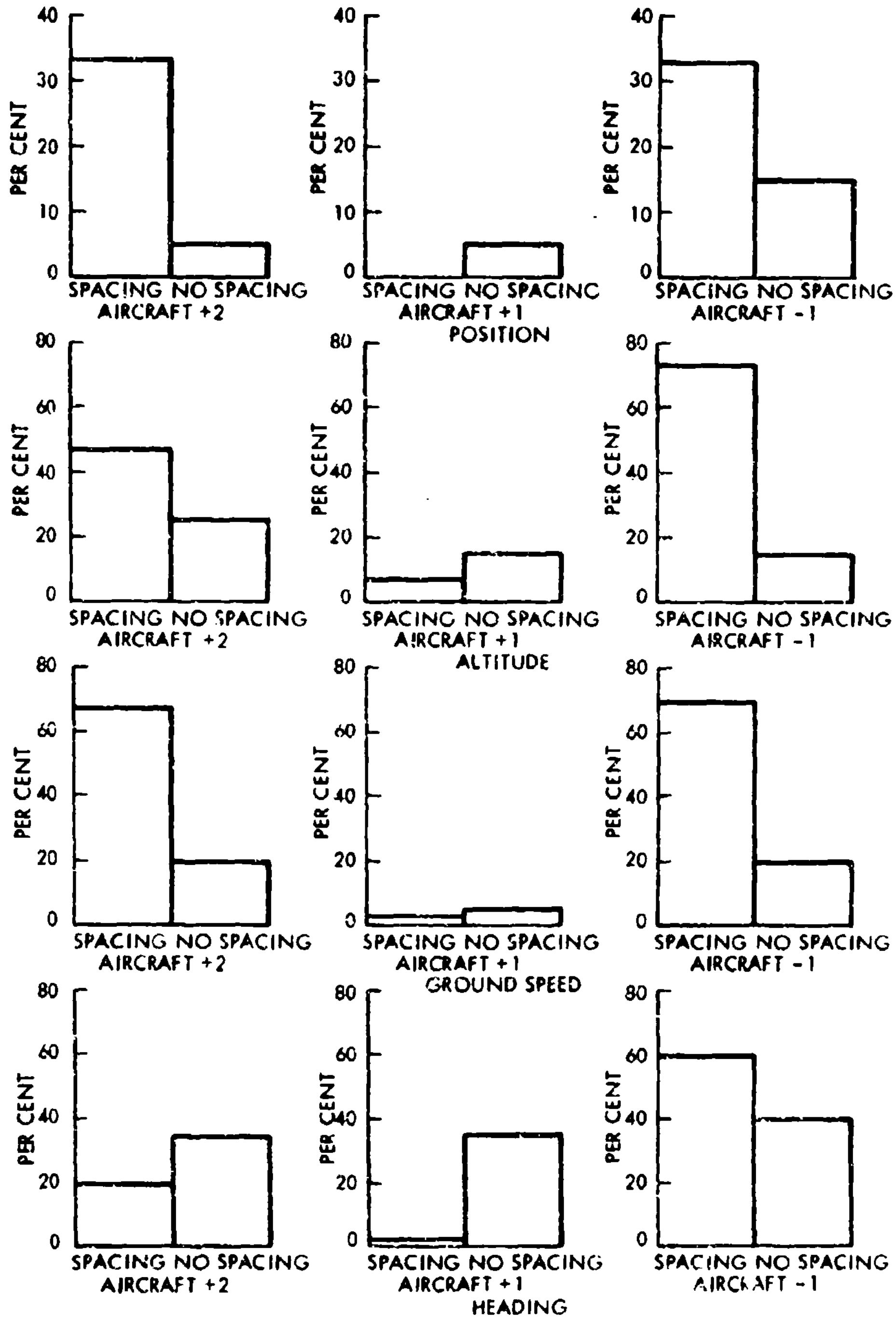


Figure 3.4.17

SPACING TASK TREATMENT: NULL RESPONSES OF INFORMATION COMPONENTS (PRE-ANALYSIS)

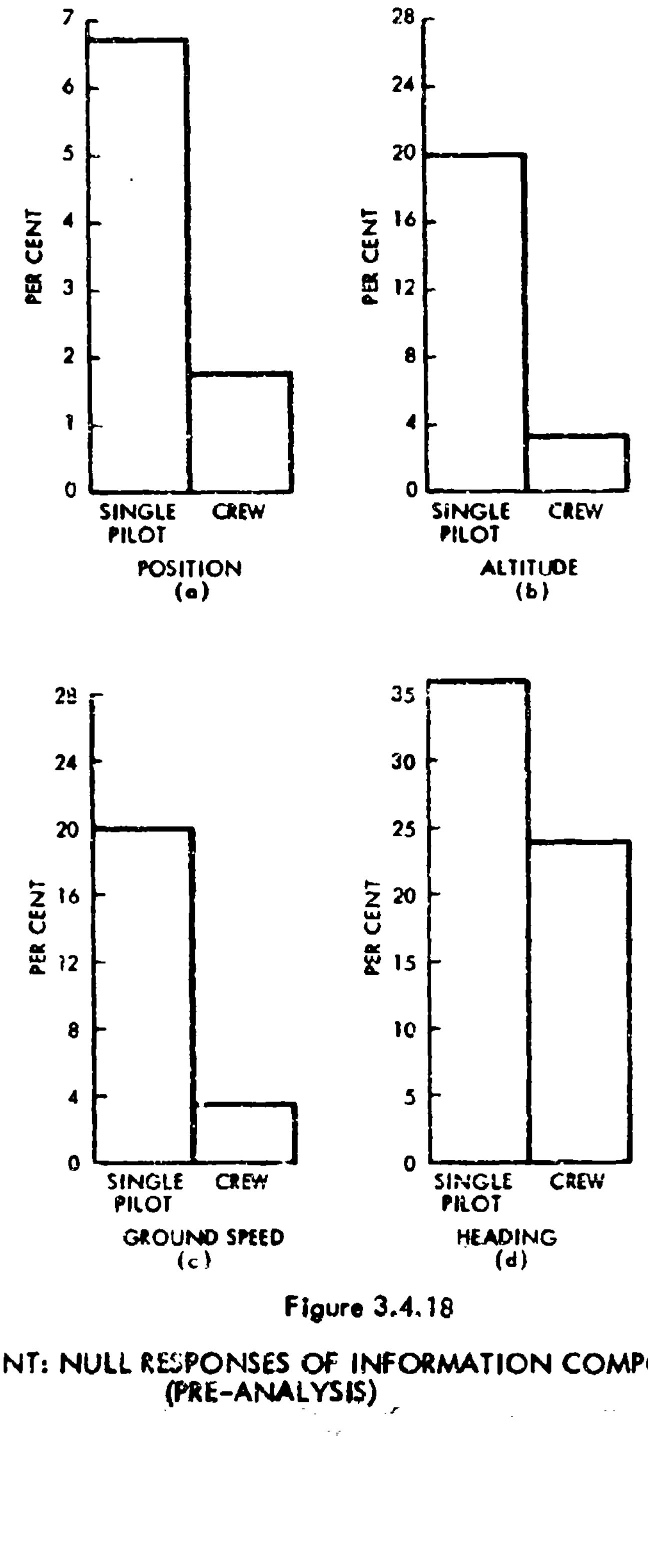


Figure 3.4.18

CREW TREATMENT: NULL RESPONSES OF INFORMATION COMPONENTS (PRE-ANALYSIS) pooling of situations and/or aircraft was guided by the situation and aircraft null hypothesis results. Histograms of aircraft Identification Scores and Landing Sequence Scores for single pilots and crews are shown in Figures 3.4.19 and 3.4.20.

3.4.2 Conflict Detection

Main effects of reaction time data were pooled whenever no statistical signficance was shown. The results of this pooling is presented in Tables 3.4.1 and 3.4.2. Table 3.4.1 is a summary of pooled update reaction time statistics (single pilot, non-alarm subjects only.) Table 3.4.2 is a summary of pooled crew, alarm, and update rate reaction time data.

3.4.3 Pilot Opinion Questionnaire

Pilot opinion questionnaires here refer to those submitted by participants in the Phase II set of tests only. The pilot opinion questionnaire consists of two sections. The questions of the first section pertain to pilot awareness, workload, and conflict detection and resolution problems. The questions of the second section deal with simulation realism.

Histograms of the pilot responses to questions of the first section are presented in Figures 3.4.21 thru 3.4.26. The histograms have been broken down into main effect categories (single pilot, crew; alarm, no alarm; one second update rate, four second update rate) for comparison. In addition, the subjects in all the treatment categories for a given question

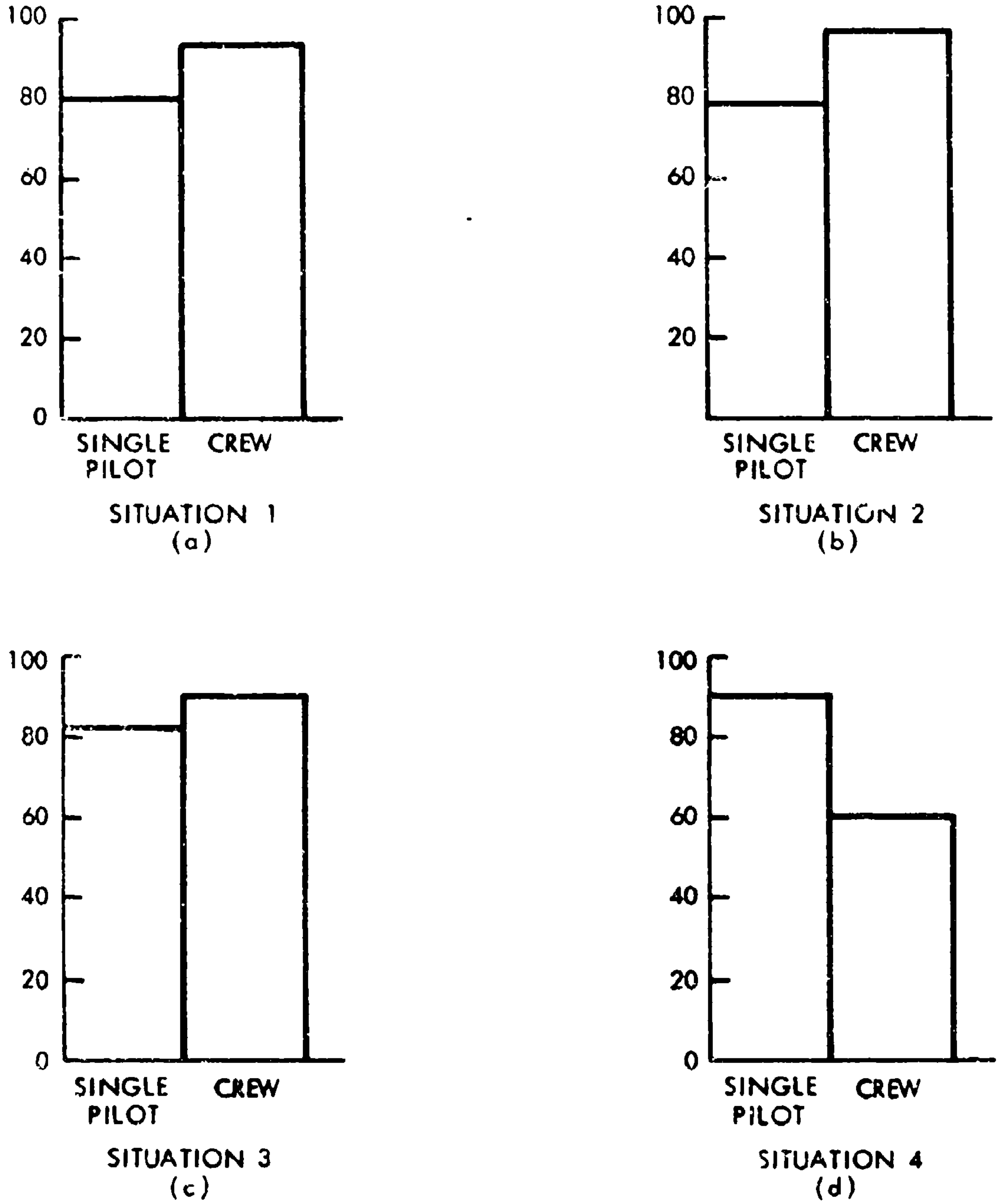


Figure 3.4.19

CREW TREATMENT: IDENTIFICATION SCORES
(POST-ANALYSIS)

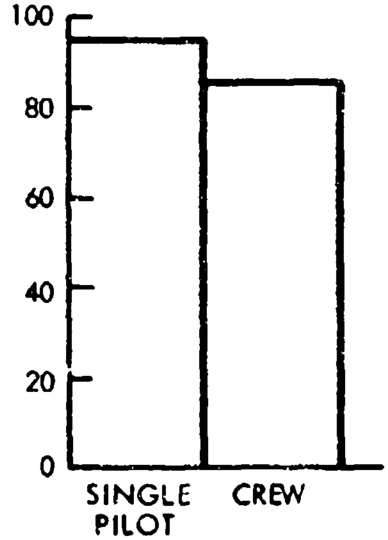


Figure 3.4.20

CREW TREATMENT: LANDING SEQUENCE SCORES
(POST-ANALYSIS)

TABLE 3.4.1

Pooled Update Rate Reaction Time Statistics

CREW TREATMENT	SINGLE PILOTS		
ALARM TREATMENT	NO ALARM		
UPDATE RATE TREATMENT	ONE AND FOUR SECONDS		
STATISTICAL PARAMETER	No.	M	S.D.
SITUATION 6	7	11.3	2.7
SITUATION 7	8	11.3	2.3

No. ■ Number of Data Points
M ■ Mean
S.D. ■ Standard Deviation

TABLE 3.4.2

Pooled Reaction Time Statistics

CREW TREATMENT	SINGLE PILOTS AND CREWS		
SITUATION TREATMENT	SITUATION 6		SITUATION 7
UPDATE RATE TREATMENT	ONE AND FOUR SECONDS		ONE AND FOUR SECONDS
ALARM TREATMENT	ALARM	NO ALARM	ALARM AND NO ALARM
NUMBER OF DATA POINTS	4	9	17
MEAN	5.2	11.7	11.9
STANDARD DEVIATION	0.4	3.1	2.3

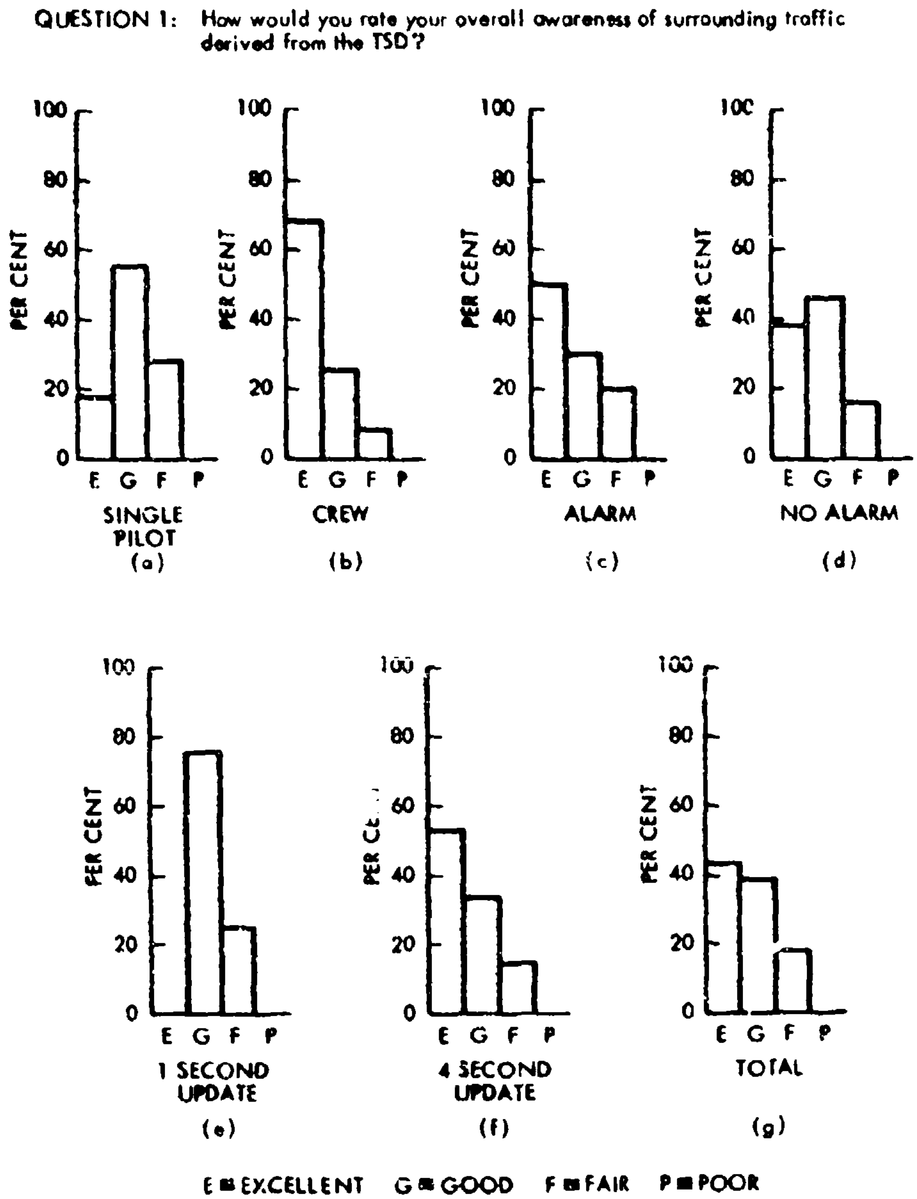
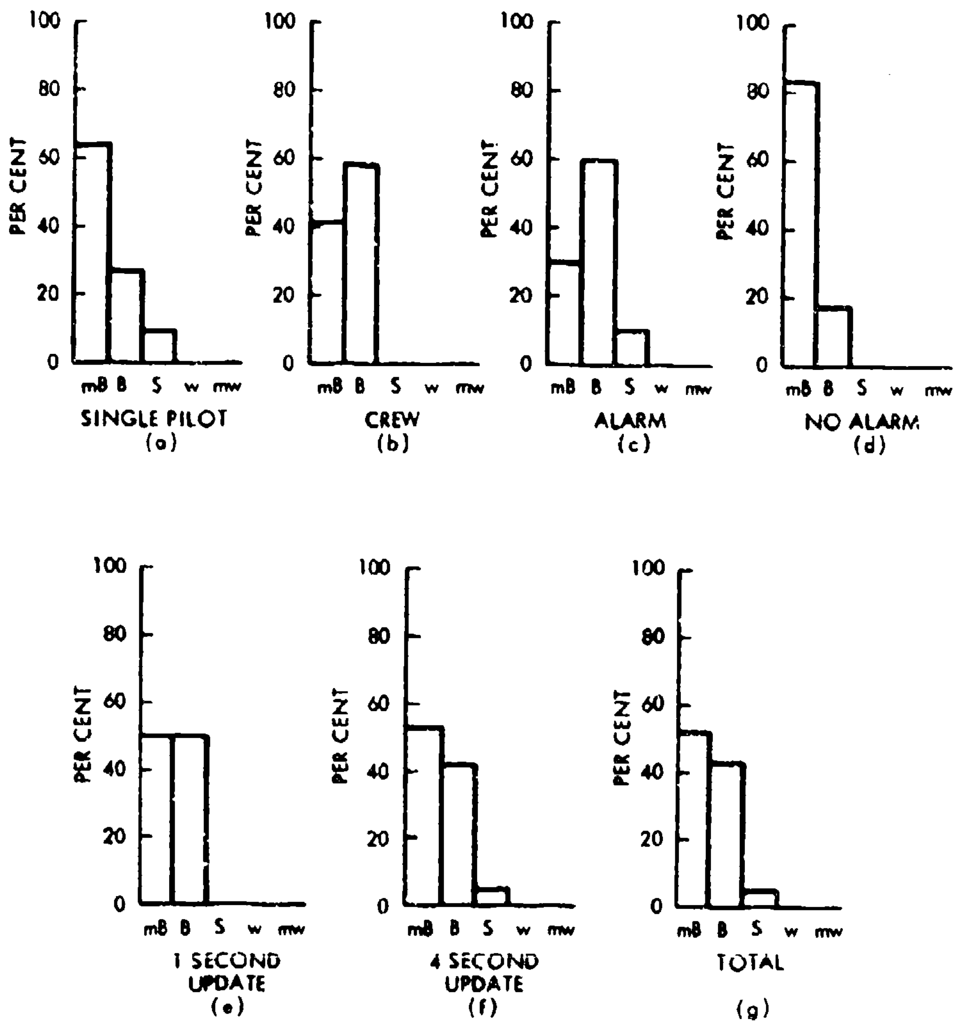


Figure 3.4.21

PILOT OPINION QUESTIONNAIRE RESULTS FOR QUESTION 1

QUESTION 2: How does this compare with your awareness in the current air traffic control system under instrument meteorological conditions?

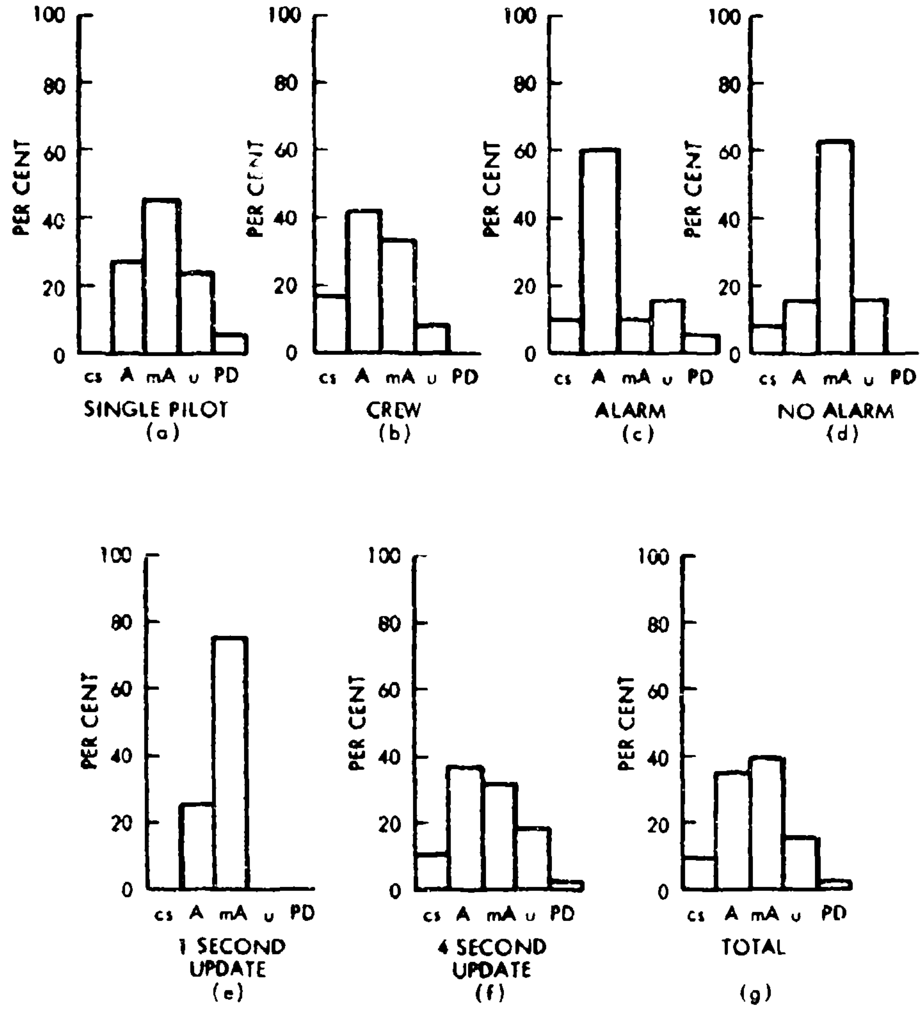


mB = MUCH BETTER B = BETTER S = SAME w = WORSE mw = M CH WORSE

Figure 3.4.22

PILOT OPINION QUESTIONNAIRE RESULTS FOR QUESTION 2

QUESTION 3: How would you rate the safety associated with closely spaced, IFR operations using the TSD?

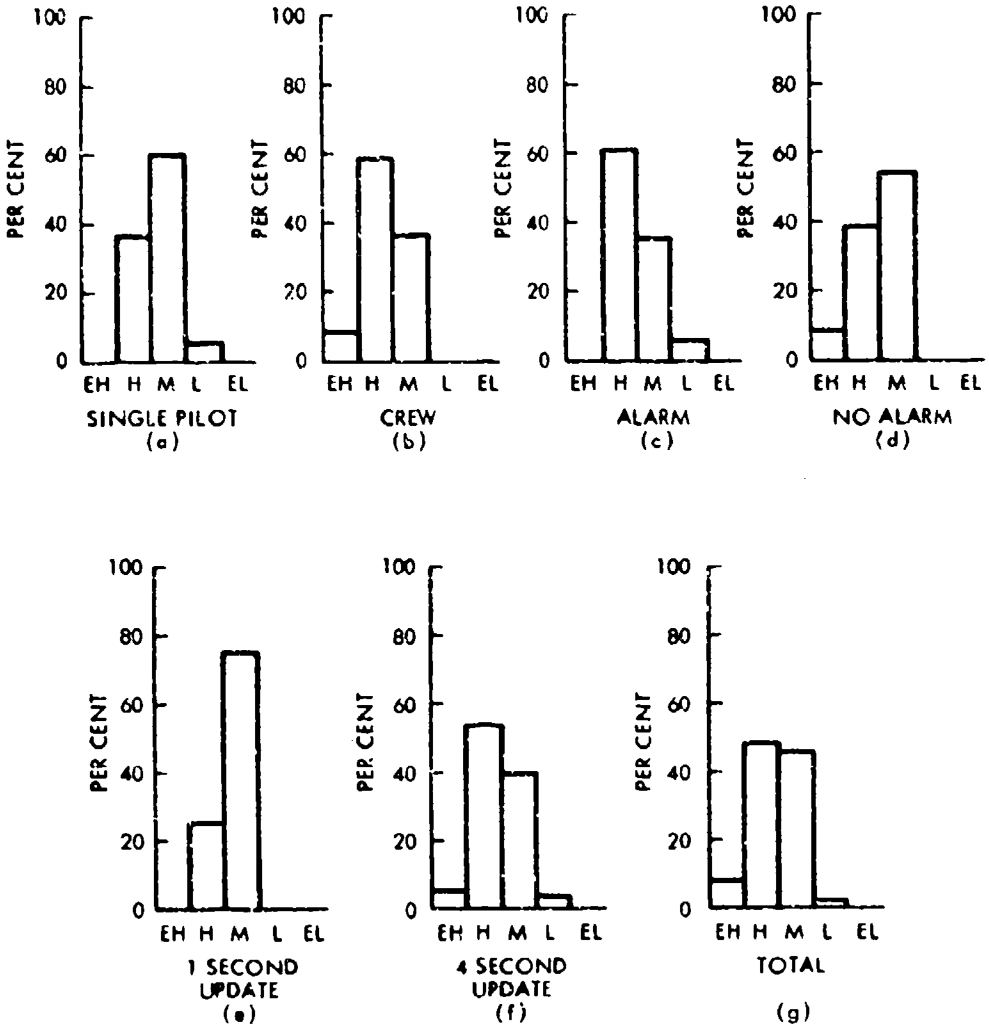


cs = COMPLETELY SAFE A = ACCEPTABLE mA = MARGINALLY ACCEPTABLE
u = UNACCEPTABLE PD = POSITIVELY DANGEROUS

Figure 3.4.23

PILOT OPINION QUESTIONNAIRE RESULTS FOR QUESTION 3

QUESTION 4: How would you rate your confidence in being able to detect potential conflicts during closely spaced, parallel, IFR operations using the TSD?



EH = EXTREMELY HIGH H = HIGH M = MODERATE L = LOW
EL = EXTREMELY LOW

Figure 3.4.24

PILOT OPINION QUESTIONNAIRE RESULTS FOR QUESTION 4

QUESTION 5: How would you rate your confidence in being able to resolve potential conflicts during closely spaced, parallel IFR operations utilizing the TSD?

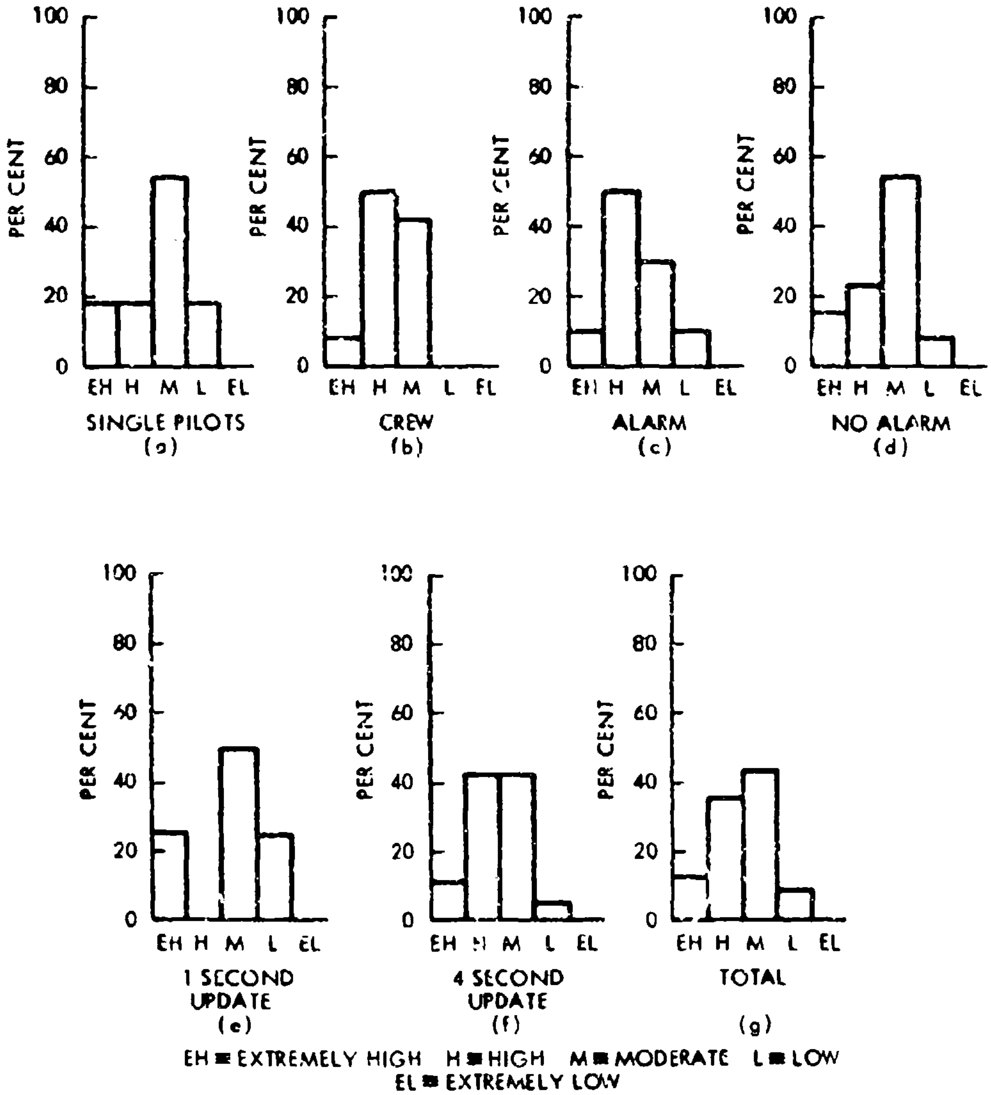


Figure 3.4.25

PILOT OPINION QUESTIONNAIRE RESULTS FOR QUESTION 5

QUESTION 6: What is your evaluation of the overall workload level associated with the simulator tasks utilizing the TSD compared to the normal cockpit duties you encounter?

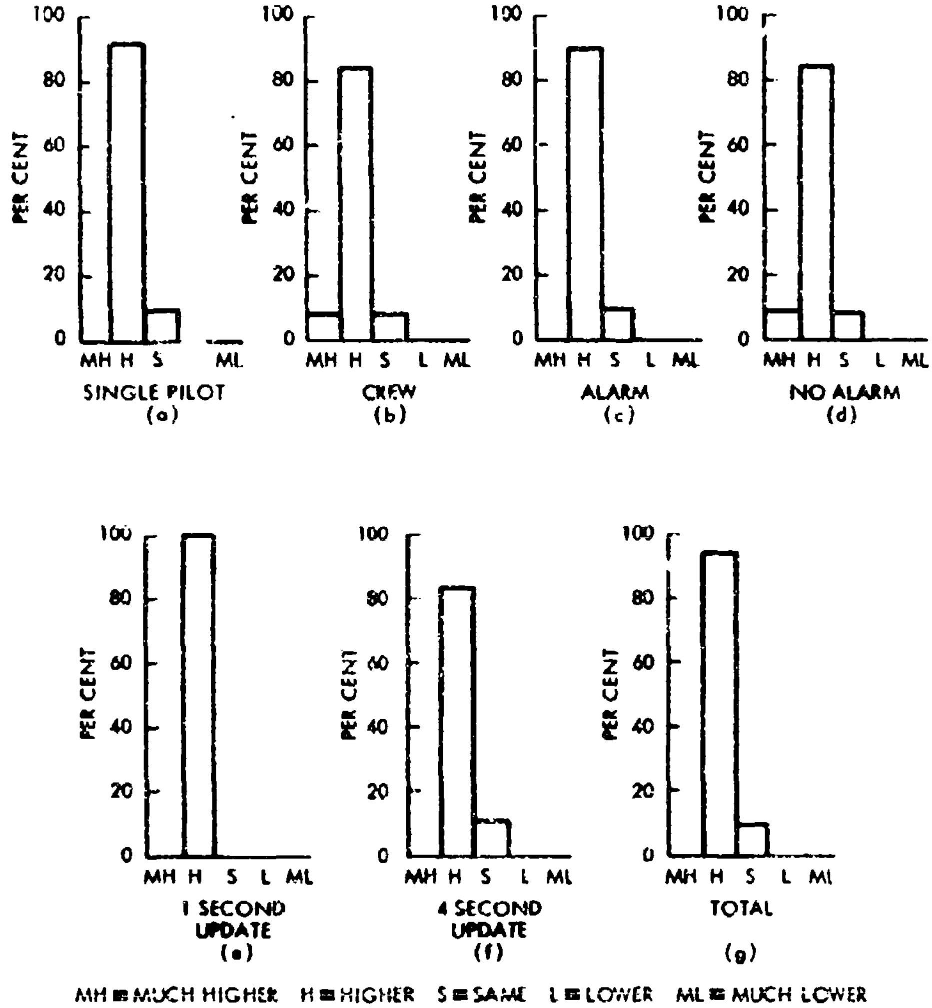


Figure 3.4.26

PILOT OPINION QUESTIONNAIRE RESULTS FOR QUESTION 6

were pooled to give a total distribution of responses.

Figure 3.4.27 presents histograms of the pilot's responses to the second section of the questionnaire. Each category has two parts. The first part shows the distribution of responses by degree of realism; i.e., excellent, good, fair and poor. The second part indicates the pilot's thoughts on whether or not realism was sufficient for the goals of the present research. Pilot's responses here were either adequate or inadequate.

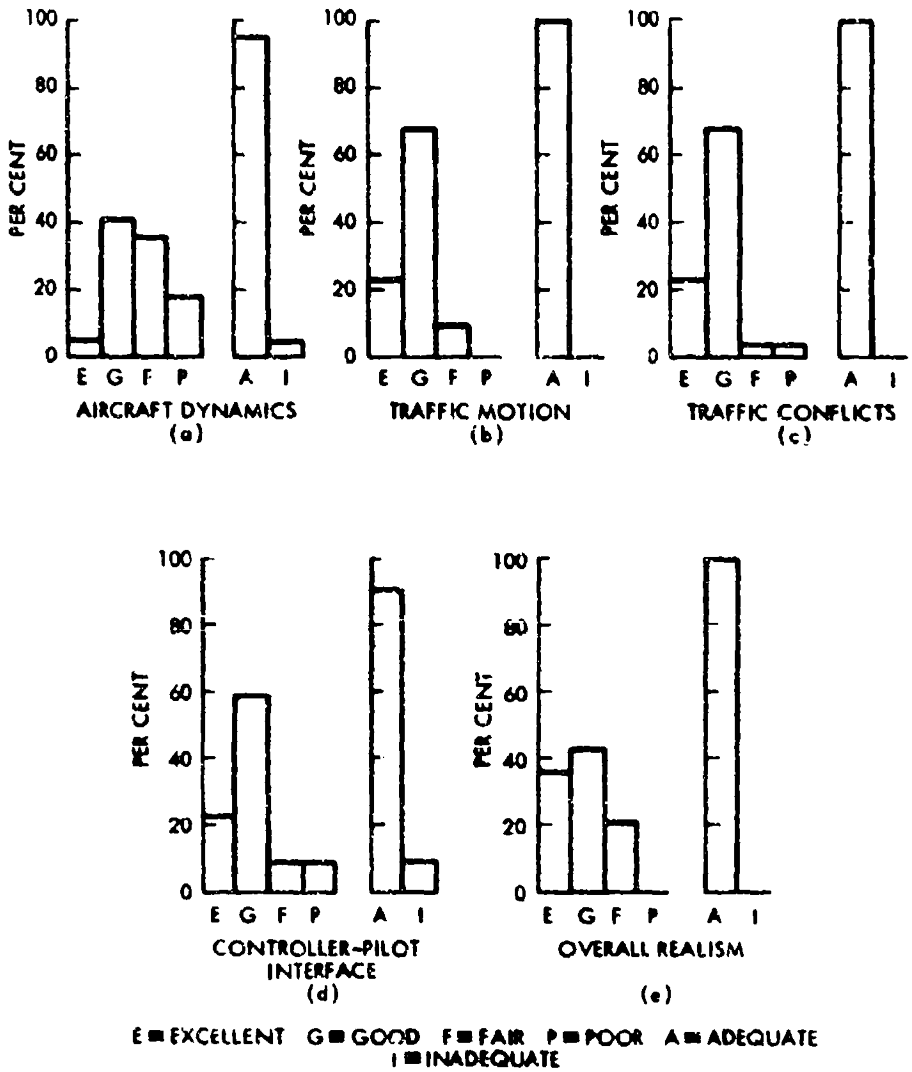


Figure 3.4.27

RESULTS OF SIMULATION REALISM QUESTIONNAIRE

CHAPTER 4

DISCUSSION

This chapter presents a discussion of the results summarized in the Tables and Figures of Chapter 3. Because real-world situations were simulated, cause and effect relationships are much more complicated than would be the case with carefully contrived abstracted experiments.

4.1 Information Components

The party-line communication channel allows the pilot to monitor air traffic control (ATC) radio transmissions. These transmissions give the pilot specific indications of the information components of surrounding aircraft. A communication concerning a given information component of any given aircraft occurs rather infrequently. In fact, such a transmission usually occurs only when a change in that component is desired by the approach controller. Moreover, if the pilot has no other means of monitoring traffic motion, he must assume that each aircraft is following instructions. This assumption is not always justified. Misinterpretations and blunders do occur.

The ATSD, on the other hand, provides the pilot with a nearly continuous source of information, but, by its nature, requires that the pilot specifically seek out a desired information component. In addition, the display acts as an auxiliary memory, thus eliminating the necessity of memorizing information components.

A summary of the results discussed at length in the following sections is offered here.

Estimation of information components tended to be situation dependent. In most cases, there was either a significant main effect or interaction involving the aircraft sequence treatment. Altitude, ground speed, and heading components tended to be estimated fairly accurately when an aircraft was in close proximity to the subject and when these quantities were not changing. Substantial mean errors and standard deviations usually showed up in information components of aircraft that were remote with respect to the subject or in components that either were changing at the time of the quiz or had changed just prior to the quiz. In addition, the heading component was estimated quite poorly when the aircraft in question was off the STAR.

Very few statistically significant major treatment results were observed in the results of the analysis of variance. The contingency tests performed on the null reresponses tended to be a more sensitive measure of differences due to these treatments.

Three major results came from the null response analysis. First, except for the ground speed component, there is no statistical significance in the differences in missing responses due to ATC treatment. Second, there is a significant difference by aircraft in the ability of pilots with the ATSD and an in-trail spacing task to respond to most information components. The target aircraft (+1) always has a very low or zero percentage of null responses while the other aircraft have high percentages of null responses. In addition, there is, in most cases, a significant effect of spacing task treatment on null responses. Third, in many cases, the two-man crews do significantly better in responding to components than do single pilots.

4.1.1 ATC Treatment

The ATC major treatment variable compares the pilot's ability to acquire traffic information via two different sensory modalities. The "Today's System" treatment provides an auditory information source consisting of party-line ATC transmissions while the "Future System" treatment provides a visual source of information via the ATSD. Radar vectors concerning the subject's own aircraft were given when the display was employed, but, unlike the party-line cases, instructions to other aircraft could not be monitored.

Radio transmissions consisted of initial contact messages (at or near one of the holding fixes), approach clearances (at ILS turn on), and landing clearances (between outer and middle markers). In addition, commanded heading, altitude and speed changes were issued.

4.1.1.1 Position Error

The pre-analysis composite means of Figure 3.2.1 and the three way analysis of variance results of Table 3.3.1

indicate a significant effect due to both ATC treatment and situation. Table 3.3.1 also indicates an interaction between the three main effects (ATC treatment, situation and aircraft).

This second order interaction means that the effect of these three treatments on position error are not independent. Consequently, a second level breakdown (by situation and aircraft) analysis was performed. This analysis, in essence, compared the ATC treatment effect aircraft by aircraft in each situation.

According to Table 3.3.6 the only aircraft to show significant ATC effects were the +1 aircraft of situation 2 and the -1 aircraft of situation 3. In both cases, the mean error was less when the ATSD was used.

Just prior to the stop-action quiz in situation 2, the -1 and +2 aircraft had received messages containing information placing them near a holding fix and the ILS turn on point (gate) respectively. Such specific position information concerning the subject and +1 aircraft was not as current. This accounts very well for the observed trend in the Today's System date. Because the with-display subjects could monitor aircraft positions as frequently as other duties permitted, they did, on the average, considerably better in estimating position for their own craft and the one just ahead on the landing sequence. The only statistically significantdifference, however, was in the estimates for the

+1 aircraft.

In situation 3, the -1 aircraft was coming around on a missed approach and therefore was not flying on any of the route segments. Radio transmissions to that craft consisted of a resequencing message and a heading change command which the aircraft's pilot misunderstood. The with-display subjects could, of course, monitor the -1 aircraft with the ATSD and thus were better informed as to its position.

The +1 and +2 aircraft were both on the ILS and had recently received transmissions with specific position information. The without-display subjects, therefore, did an equally good job as the with-display subjects in estimating the positions of these aircraft.

Just prior to the stop-action quiz, the subject's aircraft had received a heading change command from the approach controller. This gave them a reference to the route structure on which to base a position estimate.

Figure 3.4.1 indicates the best estimate under this treatment as to a pilot's information concerning aircraft position. Although the magnitude of the mean errors may seem high, it should be remembered that the scale of the quiz map was only one inch for every four miles.

As can be seen in Table 3.3.14, the aircraft hypothesis for null responses was verified, while the situation hypothesis was refuted for the position component under the Today's

System treatment. The percent null response histograms of Figure 3.4.16 (a) are therefore broken down by situation.

In situation 2, about 7.4 percent of the aircraft went undetected without the ATSD while the with-display pilots did not miss any. This difference, however, was found to be statistically insignificant as can be seen in Table 3.3.17.

In situation 3, ten percent of the aircraft failed to be reported without the ATSD, while eleven percent were not detected when the ATSD was employed. Again, this difference was found to be insignificant.

4.1.1.2 Spacing Error

Figure 3.2.2 indicates that, at least in situation 2, pilots with the display had a much better idea of how far away the other aircraft were. It can be seen that when the pilots operated in the present day display/communications environment their average spacing error for the +1 aircraft was 7.4 nautical miles. For those pilots operating in a proposed future system employing the ATSD, the average spacing error for this same aircraft decreased to .4 nautical miles. Such major improvements in spacing estimates for situation 3 are not evident.

The results of the analysis of variance presented in Tables 3.3.1 and 3.3.6 substantiate the superiority of the visual information source in estimating aircraft separation. It was found that accuracy in spacing estimates is situation dependent and that in situation 2 significant results were obtained for both the +1 and +2 aircraft. No statistical significance was observed for situation 3.

Figure 3.4.2 indicates the post-analysis statistics for spacing error under the ATC treatment variable. Both positive and negative errors are shown. A negative error means that, on the average, subjects thought the aircraft was farther away than it really was. All but one of the aircraft responses are negative. This suggests a false sense of security on the part of many pilots.

The spacing task null response data are identical to that of position error and is discussed in the section 4.1.1.1.

4.1.1.3 Altitude Error

At face value, Figure 3.2.3 seems to indicate a substantial difference in altitude error for situation 2 (especially for aircraft +1). As can be seen from Tables 3.3.1 and 3.3.6, analysis does not bear this out.

The analysis does indicate, however, that there is a significant effect due to the particular aircraft being considered. This seems plausible. Except for one distortion (aircraft +1 of situation 2), the magnitudes of errors seem to follow the same trend and considerable differences in these magnitudes for different aircraft can be observed.

Using Figure 3.4.3 as a reference, the aircraft effect can be explained. The +2 aircraft is flying straight and level at 1900 feet on the ILS. Estimates for this aircraft are quite good. This seems reasonable because the altitude for this aircraft has been constant for a considerable length of time preceding the stop-action quiz.

The +1 aircraft has just transitioned from 6000 feet to 2000 feet at 2000 feet per minute descent rate. The large mean error (-1200 feet) and substantial standard deviation ($\pm$1800 feet) indicate that the rate of change of this component makes it difficult to monitor efficiently.

The -1 aircraft, like the +1 aircraft, is flying straight and level, though at 6000 feet rather than 1900 feet. If the altitude component of this aircraft were being monitored as closely as the +2 aircraft, it would be expected that mean errors and standard deviations for the two aircraft would be similiar. This does not appear to be the case. This lack of attention in altitude estimations for aircraft behind the subject in the landing sequence is indicative of a general trend of degradated responses for the -1 aircraft for many information components.

Situation 3 seems to exhibit just the opposite trends from those that are observed in situation 2. The -1 aircraft is estimated rather well. In this scenario, the -1 aircraft has performed a missed approach procedure and is resequencing in the traffic flow. This fact is enough to distinguish this aircraft as an abnormal situation and accord it special attention. The fact that it is flying at constant altitude (1900 feet) also makes it easier to respond correctly.

At the stop-action quiz, the +1 aircraft is at about the same place as the +2 aircraft of situation 2. Consequently, it is not surprising that the responses to these aircraft are quite similar.

Pilots seem to have difficulty in estimating the altitude of the +2 aircraft in situation 3, although not nearly as much trouble as the situation 2, +1 and -1 aircraft. It happens that the +2 aircraft in this case is past the outer marker and is descending at 800 feet per minute on final approach.

It appears that the pilot's ability, with either information source, to monitor altitude is aircraft dependent and, further, is in some way related to the rate of change of altitude. In addition, it seems probable that aircraft behind the subject's in the landing sequence are, to a great extent, ignored unless some special situation (e.g., a go-around) attracts the pilot's attention.

That pilots utilizing only the party-line communication channel as an information source fail to make altitude component responses a large percentage of the time is evidenced by Figure 3.4.16 (b). Null responses with the display is cut to just over half of what it is without the display (35 percent verses 19 percent). The statistical significance of this improvement is borderline as shown in Table 3.3.17.

If a real difference in null responses does exist, it would not be very surprising. Specific altitude information for other aircraft comes rather infrequently to the pilots depending on party-line communication. For the most part, subjects must use their knowledge of the route structure and a bit of extrapolation or interpolation to make estimates. Many pilots evidently do not feel confident in making such an estimate and therefore do not respond.

Although the two information source treatments (present or future ATC display/communication format) require different monitoring techniques and supply information in different ways and at different frequencies, the result, insomuch as altitude is concerned, appears to be very nearly the same.

4.1.1.4 Ground Speed Error

Table 3.3.1 indicates that there is a significant effect on ground speed error due to ATC and aircraft treatment, but that a number of interactions necessitate a second level breakdown analysis by situation and aircraft. The results of Table 3.3.6 show that the only significant differences for the ground speed statistics presented in Figure 3.2.5 are in the +1 aircraft of both situations. In both cases, the magnitude of the mean error associated with the ATSD information source is considerably less than that provided by the party-line communication channel.

Ground speed error responses follow the same basic trends that were observed for altitude error. Those aircraft that were in the process of, or had just finished, reducing speed were estimated less accurately than those that were stable. This fact becomes obvious if the results presented in Figure 3.2.5 are explained in light of the ground speed histories of the various aircraft just prior to the stop-action quiz.

In situation 2, the speed of the -1 aircraft has been constant at 200 knots throughout the entire pre-quiz portion of the scenario. It is not surprising, therefore, that both information sources allow the pilot to estimate this component with great accuracy. In fact, the combined mean error for this aircraft is approximately zero.

The +1 aircraft is in transition from 200 knots to 160 knots. The subjects did a much poorer job in estimating the speed in this case, but those with the display did significantly better than those without it.

The +2 aircraft had been stable at 160 knots for a reasonable length of time prior to the stop-action quiz, but had reduced from an original speed of 200 knots earlier in the simulation. Both with and without display subjects did reasonably well in estimating this variable. The combined mean error was only about 3 knots.

In situation 3, the -1 aircraft is in the process of transitioning from 200 knots to 160 knots. At the time of the stop-action quiz, it has achieved the midpoint in its speed reduction with a ground speed of 180 knots. Pilots flying the simulator under both ATC treatments did a relatively poor job in estimating ground speed for this aircraft. The average error was not too large (8 knots), but the spread of the responses had a standard deviation of about 24 knots.

Significantly different results were observed for the +1 aircraft. The without-display subjects had a mean error of -24 knots, while the with-display subjects had a mean error of only about -6 knots. This aircraft had a constant speed of 160 knots for a reasonable amount of time prior to the stop-action quiz, but had reduced from an original speed of 200 knots earlier in the simulation.

The +2 aircraft had just completed a speed reduction from 160 knots to 131 knots at quiz time. The mean error in this case was almost -8 knots with a standard deviation of about 13 knots. This represents a slightly less accurate estimation than for aircraft that had been at constant speed for longer periods of time prior to the stop-action quiz. The dispersion of the data, as indicated by the standard deviation, is comparable to that obtained with other aircraft at a stable speed.

As can be seen from Figure 3.4.16 (c) and Table 3.3.17, the number of missing (null) responses is significantly higher for the party-line communication with no ATSD treatment. In fact, subjects in this treatment failed to respond to the ground speed component about 43 percent of the time.

Since the percent null responses for the discrete communications with ATSD treatment was only about 5 percent, it would seem that pilots in the present ATC display/communications system have a considerably more difficult time in picking up ground speed information than those in the proposed future system with an ATSD.

4.1.1.5 Heading Error

Before discussing the heading error components in detail, a comparison should be made between the information presentation formats of the two major treatment variables.

The party-line communications channel provides the pilot with specific heading information in that heading change commands are a part of the radio transmissions. The ATSD does not provide a specific heading readout. When the display is employed, the pilots must obtain heading information from the track history provided by the tracer dots as referenced to known courses on the display.

Table 3.3.1 indicates that the only significant main effect for heading error is due to aircraft treatment. In addition, all first order interactions are significant. This last fact necessitated a second level breakdown analysis (by situation and aircraft). This breakdown analysis, presented in Table 3.3.6, shows that significance is achieved by only the -1 aircraft of situation 3. A look at Figure 3.2.4 reveals that there is a difference of about 108 degress in the magnitude of the means of the responses for this component. The party-line communications (aural information source) had the better average estimate.

In situation 2, the -1 aircraft has been on a constant heading of 058 degrees from the beginning of the simulation until the stop-action quiz point. As can be seen from Figure 3.4.5 (a), both the mean and the standard deviation of these estimates are quite small (-8 and 8 degrees respectively).

The +1 aircraft has just turned from a heading of 330 degrees to 360 degrees at the quiz stop point. The mean error for this aircraft is very close to zero, but the standard deviation in the estimates is quite high. It is interesting to note that the standard deviation is approximately the same as the difference in the heading change (34 degrees vs. 30 degrees). It seems obvious that a number of pilots had not yet become aware of the heading change at the time of the stop-action quiz.

The +2 aircraft was well established on the ILS at the time of the quiz. It is therefore surprising that the mean error and standard deviation in the estimates for this aircraft are so large. The pre-analysis composite graph of Figure 3.2.4 does indicate a reasonably good mean error for the with-display subjects, but the differences with respect to the without-display responses fails to be statistically significant.

As previously noted, in situation 3 the -1 aircraft is executing a go-around procedure and is in the process of re-entering the landing sequence. Subsequent to the quiz stop point, a heading change had been ordered for this aircraft which the pilot read back incorrectly. The ordered heading change was from 220 degrees to 255 degrees, while the actual change was from 220 degrees to 355 degrees.

The mean error in the responses for this aircraft under the party-line communication treatment was quite good (about -5 degrees). This is contrasted to the significantly poorer estimates made by the ATSD subjects (average error of 112 degrees). The with-display subjects did, however, have a much smaller dispersion in their estimates as measured by the standard deviations (24 degrees and 80 degrees respectively).

The very large standard deviation associated with the nondisplay treatment can be explained by the fact that some of the subjects in this treatment used the commanded heading (255 degrees), while others used the read-back heading (355 degrees).

It is believed that because the -1 aircraft was off the nominal route structure and therefore lacked a course reference, the pilots in the with-display treatment had a more difficult time in interpreting heading information. It is worth noting that the mean error for these pilots is reasonably close to the difference in the heading change (112 degrees vs. 135 degrees). When the fact that the turn had not yet been completed at the quiz stop time is considered, these numbers become even closer.

The responses under both ATC treatments for the +1 and +2 aircraft of situation 3 are very nearly the same. The mean error in both cases is close to zero and the standard deviations are quite small. The high degree of estimation accuracy can be accounted for by the fact that both aircraft are well established on the ILS course by the time the stop-action quiz occurs.

The percentage of null responses for the with-display subjects (37 percent) was higher than that for the without display subjects (21 percent), although this difference falls just short of statistical significance. Null responses for both ATC treatments did not show a dependence on situation or aircraft.

The high percentage of heading null responses for the ATSD subjects is indicative of the difficulty in obtaining such information from the display.

4.1.2 Spacing Task Treatment

The inclusion of the in-trail spacing task requires that the pilot use the ATSD to acquire and maintain a specified separation with respect to an aircraft. The desired separation was either three or four miles depending on the scenario. The subjects started the simulations at some initial separation, which was greater than the desired separation, and had to merge with the traffic flow and acquire the proper spacing interval.

The subjects with the spacing task were not given radar vectors, but were asked to use the STAR displayed as a nominal flight path. They did have, however, a great deal of flexibility in deviating from the STAR to achieve the in-trail spacing.

The data used for the nonspacing task condition are the same as were used for the with-ATSD (future system) treatment condition. The with-spacing task data was gathered during Howell's experiments.

The reason for including the spacing task was to see how accurately subjects could perform this task. It is thought that with the aid of the ATSD pilots can decrease aircraft separations during IMC and achieve the higher landing capacities typical of VMC operations.

4.1.2.1 Position Error

Figure 3.2.6 shows a close agreement between the magnitudes of the mean errors for both the with and without spacing task treatments. The only main effect found to be significant was the aircraft treatment.

As indicated in Table 3.3.2, neither the spacing task treatment nor the situation are significant. It might be expected that since the with-spacing subjects were forced to monitor the +1 aircraft, their estimates for that aircraft would be better than their estimates for the other aircraft.

According to the pre-analysis data, the expectation concerning a better estimate for the +1 aircraft for the with-spacing subjects on the +1 also seems to be true in both scenarios.

These expectations do not, however, hold up under statistical analysis. The analysis of variance indicates no difference between spacing treatments and, although an aircraft effect is indicated, Figure 3.4.6 says that the +2 aircraft is estimated best. In fact, this graph indicates that aircraft position estimation is a monotonically decreasing function of relative landing sequence.

The fact that the position of the -1 aircraft was estimated the least accurately is not surprising in light of previous observations that pilots seem to pay less attention to this aircraft.

Because Table 3.3.15 (b) indicates that the aircraft hypothesis for null responses for position error is very close to being refuted, these responses are broken down by aircraft. The results are presented in Figure 3.4.17.

The null responses seem to conform to expectations. The with-spacing subjects have high percentages of null responses for the +2 and -1 aircraft (33 percent each) and no null responses for the +1 aircraft. The premise that the with-spacing subject focuses his attention on the +1 aircraft, thus has at least borderline statistical validity.

The relationship of with-spacing-task null responses to without-spacing-task null responses is also as expected. The with-spacing pilots have fewer (zero) null responses for the +1 aircraft than do the without-spacing pilots, but have considerably more null responses than the without-spacing subjects for the other two aircraft. This trend is, however, lacking in statistical validity in that the only case that approaches significance is the +2 aircraft and even this a a borderline case.

4.1.2.2 Altitude Error

The pre-analysis plots of the absolute value of the mean altitude error presented in Figure 3.2.8 seem to indicate very little difference due to spacing task treatment (except for the -1 aircraft of situation 2), but a strong dependence on aircraft treatment. In turn, this aircraft dependence seems to be dependent upon the situation.

These observations are substantiated by the analysis presented in Table 3.3.2. This table indicates that there is a significant aircraft effect as well as a statistically significant situation-by-aircraft interaction. The results of the first level breakdown analysis (by situation) in Table 3.3.4 indicates no significant spacing task effect, but it does indicate a statistically significant aircraft effect for both situations.

In situation 2, the +2 aircraft was flying straight and level at 1900 feet for a considerable period before the quiz stop point. Consequently, as can be seen in Figure 3.4.8, the mean error for combined estimates for this aircraft is zero. In addition, the standard deviation for this case is comparatively small (about 240 feet).

The +1 aircraft has completed its descent from 6000 feet to 2000 feet at a rate of 2000 feet per minute. This caused the average combined error to be fairly large (about -550 feet) and the dispersion of data to be great (standard deviation of about 950 feet).

The -1 aircraft was flying at constant altitude (6000 feet) from the beginning of the simulation to the time of the stop-action quiz. The fact that the mean error and standard deviation for this aircraft is quite large (700 feet and 1500 feet respectively), despite the constant altitude, is, as noted earlier, a reflection of the pilot's apparent casual interest in this aircraft.

In contrast to situation 2, the estimates for the -1 aircraft in situation 3 are quite good with a small mean error (-20 feet) and small standard deviation (50 feet). Like situation 2, this aircraft has been flying at constant altitude from the beginning of the simulation to the quiz stop point. Unlike situation 2, however, attention is called to this aircraft because it is executing a go-around and is, in fact, in the process of blundering.

At the quiz stop point, the +1 aircraft is flying at constant altitude on the ILS. Both the mean error and standard deviation statistics are reasonably good (about -80 feet and 100 feet respectively).

The +2 aircraft was in the process of descending at 800 feet per minute on final approach when the stop-action quiz occurred. Although the standard deviation in this case is about the same as the standard deviation of the situation 2, +1 aircraft (230 feet vs. 240 feet), the mean error is substantially greater (-250 feet vs. 0 feet). It should also be noted that both mean error and standard deviation for the +2 aircraft estimates of situation 3 are less than the same values for the +1 aircraft of situation 2 which was also descending, but at a greater vertical velocity.

As was the case for the ATC treatment altitude data, the accuracy of the spacing task data is dependent upon whether or not an aircraft is transitioning or just has transitioned at the quiz stop point. In addition, the magnitude of both the mean error and standard deviation of transitioning aircraft seem dependent upon the rate of change of altitude.

According to Table 3.3.15, there is a signficant effect on null responses due to aircraft sequence for the with-spacing subjects. No situation dependence is observed.

Figure 3.4.17 shows the same trend in null responses for the with-spacing subjects observed for the position component. The percentage of missing responses on aircraft +1 is quite low (8 percent), while the null response percentages for the +2 and -1 aircraft are quite high (48 and 73 percent respectively). The without-spacing null responses are more uniformly distributed with 25, 15 and 15 percent being recorded for the +2, +1 and -1 aircrafts respectively. The without-spacing task subjects, in fact, show no violation of the aircraft or situation hypothesis.

This same graph indicates that the with-spacing pilots had considerably more missing responses for the -1 and +2 aircraft than the without-spacing subjects, but had fewer null responses for the +1 aircraft. Statistically, however, the only significant difference was obtained for the -1 aircraft.

The percentage null responses for the altitude component is another example of the target aircraft attention focusing characteristic of the spacing task.

4.1.2.3 Ground Speed Error

Figure 3.2.9 shows the absolute value of the mean ground speed errors for the spacing task treatment. Although some substantial differences in these means exist (especially the +1 aircraft of situation 2 and the -1 aircraft of situation 3), statistical analysis fails to indicate a significant spacing task treatment effect (see Tables 3.3.2, 3.3.4 and 3.3.7). Table 3.3.2 does, however, indicate a significant effect due to situation treatment and a significant interaction between situation and aircraft. The first level breakdown (by situation) analysis (Table 3.3.4) yields a significant aircraft effect for situation 3, but this is confounded by the associated first order interaction. The ensuing second level breakdown analysis (Table 3.3.7) does not yield any significant effects due to spacing task on any of the aircraft.

It would seem likely that the ground speed readout of the target aircraft (+1 aircraft) would be estimated quite accurately with the spacing task since this is an important element in acquiring and maintaining separation. More specifically, the reduction of ground speed by the target aircraft provides a cue to the pilot to start his own speed reduction manuever.

This fact is certainly evident in both situations. The absolute mean errors for the +1 aircraft under the with-spacing condition are zero knots for situation 2 and 2 knots for situation 3. In both cases, these values are less than the corresponding absolute mean errors under the without-spacing task condition and better than the +2 aircraft estimates under the with-spacing task condition.

Apparently, the fact that speed changes are made rather infrequently washes out any significant difference between the accuracy of the subjects in estimating this component.

The combined graphs of Figure 3.4.9 follow the same trends observed earlier. The -1 aircraft of situation 2 has been flying at a constant ground speed of 200 knots during the interval from simulation commencement until the quiz stop point. It is apparent from the graph that this fact permitted the pilots to do an excellent job in this estimation (mean error of zero knots).

The +1 aircraft was just beginning to reduce speed from 200 knots to 160 knots. Consequently, both the mean error (14 knots) and standard deviation (16 knots) are fairly high.

The +2 aircraft, although having a stable speed of 160 knots at the time of the stop-action quiz, was just short of the point where it would begin its reduction to final approach speed (131 knots). Some pilots apparently anticipated the initiation of the speed reduction thus causing a mean error of about 7 knots and a standard deviation of about 14 knots.

In situation 3, the -1 aircraft was in the process of reducing from 200 knots to 160 knots and was, in fact, at 180 knots at the time of the quiz. The high mean error and standard deviation (12 knots and 29 knots respectively) reflect this.

The +1 aircraft was flying at a constant speed of 160 knots but was preparing to slow to 131 knots. The mean error for this aircraft is only -3 knots, but the standard deviation is 13 knots.

The +2 aircraft of situation 3 had just completed its final approach speed reduction at the quiz stop point. Again, the fact that this aircraft had made a recent change in ground speed caused a fairly large mean error (-12 knots) and standard deviation (17 knots).

The situation hypothesis for null responses was verified, but the aircraft hypothesis was very nearly refuted (see Table 3.3.15). For this reason the null response analysis was broken down by aircraft (Table 3.3.17).

Figure 3.4.17 shows that, once again, the same basic spacing task null response pattern emerges. The with-spacing percentages are very high for the +2 and -1 aircraft (68 percent in each case), while the +1 aircraft percentage under the [illegible] condition is quite small (2 percent).

The with-spacing percentages are substantially and significantly (statistically) greater than the without-spacing percentages for the +2 and -1 aircraft, while being slightly less for the +1 aircraft.

4.1.2.4 Heading Error

Table 3.3.2 indicates that all main effects and first order interactions are significant for this information component. The second level breakdown analysis, however, reveals that the +1 aircraft of situation 3 is the only aircraft exhibiting significance for the spacing task treatment (see Table 3.3.7).

The -1 aircraft of situation 2 has been flying on a constant heading of 058 degrees from the beginning of the simulation until the stop-action quiz. In Figure 3.4.10, the mean error (-7 degrees) and standard deviation (about 12 degrees) indicate that pilots do a fair job of estimating this aircraft.

The estimates of the +1 aircraft show a small mean error (about .5 degrees) but a very large standard deviation (about 25 degrees). The reason for the large data scatter seems to be that this aircraft has just barely completed a turn from a heading of 330 degrees to 360 degrees at the quiz stop point. It is interesting to note that this standard deviation is approximately equal to the pre-quiz heading change (25 degrees vs. 30 degrees). It seems apparent that some pilots had not yet detected the change.

The +2 aircraft was well established on the ILS at the time of the quiz. Consequently, it is surprising that even the fairly small dispersion of data was found.

As previously noted, the -1 aircraft of situation 3 was off the displayed STAR in the middle of resequencing into the traffic flow. In addition, it was blundering by turning too sharply to merge with the other traffic. These facts combined to make the pilot's estimates for this aircraft quite erroneous (mean of 107 degrees and standard deviation of about 46).

These results are a further indication of the relative difficulty pilots have in estimating aircraft heading when no underlying course reference is available.

The means and standard deviations of both spacing task treatment conditions for the situation 3, +1 aircraft are both quite good. They are statistically different, though, with the with-spacing condition having the better estimates. The +1 aircraft is, of course, the target aircraft for the spacing task. The situation 3, +2 aircraft is well established on the ILS and is estimated quite well.

The null response aircraft hypothesis was refuted for this information component (see Table 3.3.15 (b)). For this reason the analysis was broken down by aircraft. As can be seen in Table 3.3.15 (a), the situation hypothesis was confirmed.

The same trend in with-spacing task null responses observed in the other information components is also found here. The +2 and -1 aircraft receive a large percentage of missing responses (20 and 60 percent respectively), while the +1 (target) aircraft receives very few (2 percent).

As indicated in Table 3.3.17, there is no difference between null responses due to spacing task treatment.

4.1.3 Crew Treatment

All previous ATSD research efforts consisted of single pilot simulations. The pilot was required to perform many of the cockpit tasks normally performed by two pilots as well as monitor the ATSD. This, of course, created a high workload situation and it was thought to seriously degrade the pilot's ability to monitor the ATSD for both traffic and conflict detection information.

As a consequence of this belief, it was decided to compare the awareness associated with single pilot and two man flight crew treatments.

All data presented in this section was collected during Phase II experiments expressly for the purpose of analyzing the crew condition comparison. No spacing task was required and the party-line communication channel was not provided.

4.1.3.1 Position Error

The pre-analysis position error graphs of Figure 3.2.11 indicate a very close agreement between single pilot and crew estimates. For the most part, the crew condition tends to have more accurate responses, but according to the three way analysis of variance these differences are not statistically significant (see Table 3.3.3).

As can be seen in Table 3.3.3, no significant effect due to situation was observed, but a significant effect due to aircraft is indicated. The post-analysis plot of Figure 3.4.11 shows the previously noted ATSD position error trend, i.e., a monotonic function of relative landing sequence with the -1 and +2 aircraft being estimated the least and most accurately respectively. It is believed that the reason the higher numbered (indexed by relative landing sequence) aircraft are estimated more accurately is that there is an abundance of land-marks (ILS, outer marker, middle marker and airport) with which to reference these aircraft. Such references are less abundant elsewhere.

Neither the situation nor the aircraft hypothesis was refuted for either treatment (see Table 3.3.16). As a result, the position null response data was pooled by situation and aircraft.

The position null response histograms shown in Figure 3.4.18 (a) indicate a small percentage of missing responses under both conditions with the crew treatment having a lower percentage than the single pilots (1.8 percent verses 6.7 percent). According to Table 3.3.17, however, this difference is insignificant.

4.1.3.2 Spacing Error

Spacing error estimates, like position error estimates, follow the same trends under both crew treatments. Figure 3.2.12 indicates that there seems to be no clear cut superiority for either case. This fact is supported by the analysis results presented in Tables 3.3.3, 3.3.5 and 3.3.8 which indicate no significant effect due to crew treatment. The first table does, however, indicate a significant main effect due to situation as well as a significant situation-by-aircraft interaction. The second table indicates a significant aircraft effect for situations 1, 3 and 4.

The spacing error null responses are identical to the position error null responses which are discussed in section 4.1.3.1.

4.1.3.3 Altitude Error

Except for situation 2, altitude error also shows comparable means errors for both single pilots and two man crews (see Figure 3.2.13). This observation is supported by Table 3.3.3 which indicates no significant effect due to any of the main treatments. A situation-by-aircraft interaction in the three way analysis and an aircraft-by-crew treatment interaction (situation 1) in the first level breakdown analysis (Table 3.3.5) necessitated a second level breakdown (by situation and aircraft) analysis (Table 3.3.8). In neither the first nor second level breakdown analysis was there observed a significant effect due to crew treatment.

The -1 aircraft of situation 1 is, at the time of the stop-action quiz, flying straight and level at 2000 feet. It had descended from an altitude of 5000 feet earlier in the simulation. As can be seen from the combined graph of Figure 3.4.14 (a), the fact that the aircraft was flying at constant altitude allowed the pilots to do a good job in estimating this component. The mean error in this case was zero and the standard deviation was 20 feet.

The +1 aircraft was descending at a rate of 800 feet per minute at the quiz stop point. Although the mean error was zero, a comparatively high standard deviation was encountered (about 260 feet).

The +2 aircraft was also descending at 800 feet per minute at the time of the stop-action quiz. This aircraft had a mean error of about -160 feet and a standard deviation of about 520 feet.

At the quiz stop point in situation 2, the -1 aircraft was flying at a constant altitude of 6000 feet. Figure 3.4.14 (c) indicates a large mean error (about 420 feet) and a substantial standard deviation (about 1000 feet) in the pilot estimates for this aircraft. This aircraft is rather remote from the subject and, therefore, seems to be of less concern than other aircraft in the scenario.

The +1 aircraft has just completed its descent from 6000 feet to 2000 feet prior to the stop-action quiz. Consequently, both the mean error and standard deviation for this aircraft are fairly poor (-350 and 950 feet respectively).

The +2 aircraft at the quiz is flying straight and level at 1900 feet. As a result of this, the mean error for this craft was small (about 50 feet) and the dispersion of responses was moderate (standard deviation of about 200 feet).

The altitude estimates for the -1 aircraft of situation 3 are quite good since it is flying at constant altitude (1900 feet) and is in close proximity to the subject's aircraft at quiz time (see Figure 3.4.14 (b)).

The +1 aircraft's altitude is also constant, while the +2 aircraft is descending on final approach. Both craft are on the ILS. The constant altitude aircraft data have a fairly small mean error (about -80 feet) with a small standard deviation (about 120 feet). The descending aircraft data have approximately the same mean error (-100), but the standard deviation is considerably larger (450 feet).

In situation 4, the -1 aircraft is in the process of descending out of 5000 feet to 2000 feet at a rate of 2000 feet per minute. At the point in the simulation where the quiz was administered, this aircraft had an altitude of 3000 feet. The fact that this aircraft was transitioning in altitude and the fact that its position was fairly remote with respect to the subject's combined to seriously degrade pilot estimates of this component for this aircraft.

As can be seen in Figure 3.4.14 (d), both the mean error and standard deviation for these estimates are very large (-620 feet and 1220 feet respectively).

The +1 aircraft was, at quiz time, flying at constant altitude of 2000 feet. For this reason, the estimates in this case were comparatively good (mean of -50 feet and standard deviation of about 270 feet).

The +2 aircraft was well established on the ILS at constant altitude (1900 feet). Consequently, subjects did a very good job in their estimations for this aircraft.

As noted in both the ATC treatment and spacing task treatment discussions, the accuracy in altitude estimations, as measured by both mean error and standard deviation, is dependent upon whether or not an aircraft altitude is or recently has been changing. The constant valued altitudes tend to get estimated quite well while the transitioning aircraft tend to get estimated comparatively poorly. The only exception to this seems to be the -1 aircraft which, although having constant altitude at all situation quiz points, it estimated poorly when remotely positioned with respect to the subject (situations 2 and 4) and is estimated quite well when in close proximity (situations 1 and 3).

Neither the situation nor the aircraft hypothesis for null responses were refuted for this component. Consequently, all aircraft and situations were pooled in the crew treatment analysis (see Tables 3.3.16 and 3.3.17).

Figure 3.4.18 (b) indicates a substantial (and statistically significant) difference in percent null responses by crew treatment. The two man crews had a much lower percentage of missing altitude estimates (about 3.5 percent) than did the single pilots (20 percent).

4.1.3.4 Ground Speed Error

The pre-analysis plots of Figure 3.2.14 indicate a large difference in many of the absolute mean errors for the ground speed component. In all but one case (+1 aircraft of situation 3), the crew condition responses were better than those given by the single pilots. According to the analysis of variance, however, this difference is not significant. As can be seen in Table 3.3.3, the only significant main effect or interaction for this component is the situation treatment.

It seems that the combined statistics from both crew treatments are insensitive to whether or not an aircraft is changing speed or is stable valued, but rather are situation dependent. This is contrary to results obtained for ATC and spacing task treatments, where a dependence upon speed changes was noted.

The situation and aircraft hypotheses for null responses were verified for the ground speed component (Table 3.3.16). Consequently, all situations and aircraft were pooled in the null response analysis. The result of this analysis is presented in Table 3.3.17 and shows a significant effect of the crew treatment on missing data.

Figure 3.4.18 (c) indicates that once again the single pilots have significantly more null responses than do the two man crews (20 percent verses 3.4 percent).

4.1.3.5 Heading Error

The absolute means of the pre-analysis heading data seem to follow the same general trends for both crew treatments (see Figure 3.2.15). Although there do appear to be some fairly substantial differences in these means, Table 3.3.3 indicates that there is no significant crew treatment effect on the heading error component. This table does, however, indicate statistically significant results for both the aircraft end situation main effects as well as a situation-by-aircraft interaction.

The first level breakdown analysis also yields non-significant results for crew treatment, but does yield statistically significant results for aircraft treatment in situations 3 and 4 (see Table 3.3.5). These two situations are the conflict scenarios whose abnormalities consist of heading anomalies. The second level breakdown analysis (Table 3.3.8) again indicates no statistically significant crew treatment.

The -1 aircraft has just turned from a heading of 080 degrees to new heading of 060 degrees at the time the stop-action quiz is administered in situation 1. As a result of this, the spread in the responses for this aircraft is fairly large (standard deviation of about 16.6 degrees) even though the mean error is reasonably small (about 2.4 degrees). Apparently some pilots did not detect the heading change at the quiz point. It is interesting to note that the standard deviation of the responses and the heading change are approximately the same (16.6 degrees vs. 20 degrees).

Both the +1 and +2 aircraft are flying the ILS course and are consequently estimated quite well (Figure 3.4.15 (a)).

In situation 2, the -1 aircraft has been flying a constant heading of 058 degrees during the entire pre-quiz portion of the simulation. Pilot estimation error for this aircraft is reasonably small with a mean of -4 degrees and a standard deviation of about 7.6 degrees (Figure 3.4.15 (b).

The mean error associated with the +1 aircraft is small (about 3.5 degrees) but the standard deviation is quite large (about 26 degrees). This was probably caused by the fact that this aircraft had turned from a heading of 330 to a heading of 360 just prior to the quiz. Again, it is interesting to note that the standard deviation is approximately the same as the heading change (26 degrees vs. 30 degrees), thus indicating that some pilots had not detected the new heading prior to the stop action quiz.

The +2 aircraft was estimated quite well since it was well established on the ILS at quiz time.

In situation 3, both the +1 and +2 aircraft were well established on the ILS course and the accurate pilot heading estimations associated with these aircraft reflect this fact.

The -1 aircraft, as noted earlier, is resequencing itself in the traffic after having performed a missed approach. Its position was off the nominal STAR to the East of the field and it was turning too sharply into the traffic flow.

As can be seen in Figure 3.4.15 (c), pilots did a poor job in estimating the heading of this aircraft. The mean error of the responses was about 117 degrees while the standard deviation was 23 degrees.

Before turning, this aircraft's heading was 220 degrees. It was turning to a new heading of 355 degrees after having misinterpreted a commanded change to 255 degrees. This misinterpretation was, of course, unknown to the subjects since they were not provided with the party-line communication channel. It should be pointed out that the difference in the heading change is approximately equal to the mean heading error (135 degrees and 117 degrees respectively). When the fact that the turn had not yet been completed at the quiz point is taken into consideration, these numbers become even closer. It seems apparent that most of the subjects were unaware that the -1 aircraft was turning into the traffic.

Situation 4 is also a conflict scenario. In this case, the +1 aircraft is the blundering aircraft. This aircraft failed to make a heading change because of radio failure. It was suppose to turn from a heading of 170 degrees to a new heading of 120 degrees. Instead, it made only a partial turn to 140 degrees. In addition, if it had followed the STAR, it would have made a second turn to 060 degrees before the quiz stop point. The large mean (about 54 degrees) and standard deviation (about 55 degrees) indicate that prior to the quiz many pilots did not detect that this aircraft had deviated from the STAR.

The -1 aircraft is flying on a constant heading of 170 degrees while the +2 aircraft is flying the ILS course (035 degrees). The -1 craft is estimated with fair accuracy having a mean of about 4 degrees and a standard deviation of about 17 degrees. The +2 aircraft is estimated very well with a mean of about zero and a very small dispersion.

The heading null response aircraft and situation hypotheses were verified, thus permitting the pooling of aircraft and situations in the analysis (Table 3.3.16).

Once again, the single pilots have the higher percentage of null responses (36 percent verses 24 percent for crews), but the difference is not statistically significant (see Table 3.3.17 and Figure 3.4.18 (d)).

4.1.3.6 Identification Point Scores

As can be seen in Table 3.3.11, both the aircraft and situation hypotheses were verified for the single pilots, while only the aircraft hypothesis was verified for the two man crews. Consequently, the statistical analysis was broken down by situation, although all aircraft within a given situation were pooled.

Figure 3.4.19 indicates that for all scenarios, except situation 4, the two man crews have a higher percentage of correct responses than do the single pilots (93 vs. 80, 96 vs. 78, 90 vs. 82 and 60 vs. 90 for situations 1, 2, 3 and 4 respectively). The results of the analysis, however, indicate that the only significant results were obtained for aircraft +2, situation 4 (Table 3.3.12). This was the case in which the single pilots were superior to crews in remembering aircraft identity.

In the three situations where no significant crew treatment effect was observed, the combined statistics indicate that 84, 83 and 84 percent (situation 1, 2 and 3 respectively) of the aircraft call signs were correctly identified. In situation 4, the single pilots correctly identified 90 percent of the aircraft while the two man crews only identified 60 percent of the aircraft correctly.

4.1.3.7 Landing Sequence Point Scores

From Table 3.3.11, it can be seen that the situation and aircraft hypothesis were verified for both crew conditions. This indicates th recognition of landing sequence position is neither aircraft or situation dependent.

Table 3.3.13 indicates that both the single pilots and two man crews did an equally good job, statistically speaking, in recognizing landing sequence positions. For this analysis, the data was pooled over all situations and all aircraft.

The pre-analysis results of Figure 3.4.20 show that the subjects in the crew treatment had a higher percentage of correct responses than did the single pilots (96 percent verses 86 percent), but as indicated above, this difference was not statistically significant.

The post analysis combined statistics indicate that in the scenarios developed for this program, the pilots (both single and crew) with the ATSD and no spacing task were able to correctly state the landing sequence 91 percent of the time. Only 9 percent of the responses were either incorrect or missing.

4.2 Conflict Detection

To merely accumulate information regarding surrounding traffic elements is insufficient to attain full pilot awareness (assurance). The pilot must also be able to use this information to detect abnormal situations, especially those that pose a potential threat to his own aircraft.

The three major treatment variables (ATC, Spacing, and Crew) are examined in this section for their effect on conflict detection. In addition, the effect of information update rate and alarm algorithms on pilot reaction times are discussed. Because of the elimination of the spacing task in some tests, the relative position of the intruding aircraft varied. For this reason detection prior to the closest point of approach (CPA) is emphasized here, not miss distances.

4.2.1 Air Traffic Control Treatment

It is of interest to compare the ability of pilots to detect abnormalities using two different information sources. The no-ATSD with party-line communication channel treatment is the equivalent of today's air traffic control display/

communications environment and provides the pilot with an aural information source. The ATSD with discrete communications treatment is representative of a possible future air traffic control display/communications environment and provides the pilot with a visual information source.

Only one single runway conflict scenario (situation 3) was used to compare differences in the detection frequencies associated with each treatment. These frequencies were compared at two points, these being: at, or before, the stop-action quiz (SAQ) and at, or before, the closest point of approach (CPA).

As stated in Chapter 2, the conflict examined here consisted of a heading readback error (aural cue) and a resulting collision abeam of the subject's aircraft. At the point of nominal impact (i.e., the point where collision would occur if the intrusion went undetected) the subjects' position was on the ILS between the gate (turn on point) and the outer marker. Those subjects using the ATSD did not have the aural cue, but, of course, could monitor the situation on the display.

At the quiz stop point, the radio transmission containing the blunder cue had been issued and the intruding aircraft had begun to turn toward the ILS. The intrusion at this point was not acute.

Only one out of ten non-ATSD subjects detected the aural cue and indicated an abnormal situation on the quiz map, while four out of ten of the ATSD subjects detected the intrusion at or before the stop action quiz. As can be seen in Table 3.3.23, this difference was not statistically significant.

The without-display simulations were not continued past the quiz since subjects in this treatment had no further opportunity to ascertain that a blunder had occurred. In this case, the number of detections at the SAQ was considered to be the same as at the CPA. The with-display simulations were continued past the quiz point since subjects in this treatment could continuously monitor the situation via the ATSD. Only six out of the ten with-display subjects are considered here, however, since the others all had proximity alarms (2 mile range threshold).

Of the remaining six ATSD subjects, all detected the intruding aircraft before it would have impacted. Table 3.3.23 suggests that this represents a significant improvement over detection using the party-line communications channel.

Table 3.3.24 indicates that the empirical probability of detection P(D) at or before the closest point of approach is .1 for the present ATC system, and 1 for the possible future system utilizing an ATSD. It should be pointed out that conflict detection is situation dependent and different results could be obtained with different scenarios. At least to the extent that the blunder considered here is representative of typical conflict situations, the ATSD appears to be an asset in increasing pilot awareness and thereby aircraft safety.

4.2.2 Spacing Task Treatment

Both spacing task treatments utilize the ATSD with a discrete communication channel. The difference in the two subject groups is that one group had to perform an in-trail spacing task while the other followed radar vectors.

It was thought that the inclusion of the spacing task might detract from the pilots' ability to monitor the display for blunders. Consequently, the single pilot, no spacing task, no alarm data collected during the Phase II experiments is compared to the single pilot, no alarm, with-spacing-task data collected by Howell. Only parallel runway simulations are examined.

According to Tables 3.3.21 and 3.3.22, the situation hypothesis is verified thus indicating that, as far as detecting whether or not a blunder has occurred, there is no difference between situations 6 and 7 for either spacing task treatment. Consequently, responses for both situations are pooled in the analysis.

All subjects in both spacing task treatments detected the blunder intrusion in situation 6 before impact would have occurred. As reported by Howell, those subjects required to perform the spacing task, and therefore not require to follow the STAR, tended to displace themselves well to the outside of the ILS centerline, thus increasing the lateral separation between themselves and adjacent aircraft roughly to current standards (around 5000 feet). The pilots would acquire the ILS centerline at the outer marker and proceed with the final approach. Because of this pilot generated extra safety margin and the relatively crude method in which detection times were measured (i.e., stop watch and voice transmissions), it is not clear whether or not as many detections would have been made with 2500' lateral spacing. Table 3.3.3 in Howell's report (reference 7) indicates that at least two and possibly as many as four, of the seven subjects would not have detected the intruder before collision if they had been flying the nominal STAR course. The decreased lateral separation, however, could possibly have had the effect of increasing the pilot's vigilance and decreasing detection time.

In situation 7, only four of the seven spacing task subjects detected the crossover intrusion while all of the eight non-spacing task subjects detected the blunder. The with-spacing subjects were all flying the nominal STAR at the time of conflict, but were no longer performing a spacing task since their target aircraft had just landed.

As can be seen in Table 3.3.23, there is no statistical difference between spacing task treatments when situations 6 and 7 are combined. It is suggested, however, that judgement on this matter should be postponed until more definitive data is obtained.

Based on results of these tests, the empirical probability of detection P(D) of with-spacing subjects for both situations is .79, while the detection probability in the same circumstances for without-spacing subjects (Table 3.3.24) is 1. The combined treatment probability of detection is .85.

4.2.3 Crew Treatment

The purpose of the crew treatment conflict detection tests is to compare the difference in the ability of single pilots and two man crews in detecting blunders. In addition, it was desired to evaluate the effect of information update rate (either four seconds or one second) and some simple alarms on pilot reaction time.

4.2.3.1 Detection

All the conflict scenarios were used in this analysis (single runway: situations 3 and 4; parallel runway: situations 6 and 7). Both single runway scenarios were run with a four second update rate, while in the parallel runway scenarios, four single pilots were run with a one second update rate. The remaining subjects had a four-second update.

Table 3.3.22 indicates that there is no difference in detections under the main treatments (crew, update rate and alarm) due to situation for either the single runway or parallel runway simulations. Consequently, all single runway and all parallel runway responses for a given treatment have been pooled in the analysis.

In situation 3, four of the ten single subjects detected the intrusion at or before the stop-action quiz, while two of the five crews did the same. All subjects (single pilots and crews, alarm and non-alarm) detected the conflict before the collision would have occurred.

Four out of ten single pilots observed the blunder in situation 4 prior to the quiz, while one out of five crews did the same. All non-alarm subjects, single and crew, detected the conflict before the collision would have occurred, while all but one alarm subject (single pilot) detected the blunder at this point. The subject who failed to detect the collision did so because his alarm failed to operate. Consequently, this data point was not used in the analysis.

The analysis performed on the combined situation data indicates that there is no statistical difference between single pilots and two man crews in performing conflict detection on approaches with a single active runway.

All subjects (single and crew, alarm and non-alarm) detected the crossover intrusions in both situation 6 and situation 7. Obviously, there is no statistical difference due to crew treatment.

When the question of detection rather than reaction time is considered, there is no difference in either single or parallel runway simulations due to crew or alarm treatments.

Although a difference was expected in the crew treatment case, none was found. This may have been due to the fact that the implementation of control wheel steering allowed the pilot to monitor the display more effectively than he would have been able to with a conventional attitude control system (although not as effectively as with a true autopilot).

4.2.3.2 Reaction Time

An effort was made to pin down the pilot reaction time to the parallel runway crossover conflicts under various crew, alarm and update rate treatments. Pilot reaction time was defined to be the time delay between the start of the crossover manuever and the initiation of the constant altitude turn. In actuality, time delays were measured with respect to the time at which the emergency alarm was (or would have been) triggered; i.e., when the Tau criteria was violated. In the following discussions, this time delay is referred to as pilot response time. The nominal collision time is defined to be the time from blunder commencement to the time that a collision would occur if the subject did not perform an avoidance manuever. The parallel runway conflict scenarios used in these experiments were developed so that the nominal collision time corresponded to the time it took the intruding aircraft to cross from his ILS centerline to that of the subject. Nominal time to escape is defined as the nominal collision time minus the reaction time in question.

4.2.3.2.1 Update Rate Treatment

The effect of information update rate on pilot reaction time was examined first. These tests consisted of non-alarm simulations utilizing single pilots only.

The mean response time for the situation 6 conflict under the one second update rate condition was 6.4 seconds, while the mean time during the same scenario under the four second update rate condition was 7.2 seconds. The standard deviations associated with these were 2.4 for the one second update rate treatment and 3.5 seconds for the four second update rate condition.

In situation 7, subjects using a one second update rate had a mean response time of 1.9 seconds with a standard deviation of 2.5 seconds, while the subjects using a four second update rate had a mean response time of 0.6 seconds with a standard deviation of 2.4 seconds.

Table 3.3.18 indicates that there was no significant effect on response time due to information update rate, but that the response times are situation dependent.

The combined update rate response time data for situation 6 have a mean of 6.8 seconds and a standard deviation of 2.7 seconds. The combined data samples for situation 7 have a mean of 1.2 seconds and a standard deviation of 2.3 seconds.

It seems apparent that situation 6 is the more dangerous of the two conflict scenarios in that the average response time is longer and the time from blunder initiation to collision is shorter. For this case, the Tau threshold is violated 4.5 seconds after the intruding aircraft crosses its own ILS centerline. The average pilot takes another 6.8 seconds to notice the blunder. This means that the average pilot did not react to the conflict until 11.3 seconds after the blunder commenced. Since the nominal collision time is 19.5 seconds, this leaves the average pilot 8.2 seconds to perform an avoidance maneuver. Assuming that pilot reaction times are normally distributed about a mean value of 11.3 seconds, a delay of 14 seconds (mean plus one standard deviation of 2.7 seconds) would encompass 84% of the subject population. A 14 second reaction leaves only 4.5 seconds to perform the avoidance maneuver. In practice, of course, altitude separation would be in effect during the ILS acquisition phase and this alleviates some of the danger associated with the situation. Overshoot on acquisition is also much less likely with the ATSD.

In situation 7, the combined update rate data have a mean of 1.2 seconds and a standard deviation of 2.3 seconds. The Tau criteria in this case is violated 10.2 seconds after the beginning of the crossover blunder and the nominal collision time is 24 seconds.

It therefore takes the average pilot 11.4 seconds from the time of blunder commencement to react to the situation. This gives him 12.6 seconds to perform an escape manuever. A delay of 13.7 seconds (mean plus one standard deviation of 2.3 seconds) would encompass 84% of the subjects and would allow the pilot only a 10.3 second nominal escape time. Even with runway threshold stagger, there is very little altitude separation between aircraft making closely-spaced parallel approaches. A crossover such as that simulated in situation 7 requires a very quick reaction time to assure a safe evasive maneuver.

4.2.3.2.2 Crew and Alarm Treatments

Table 3.3.19 indicates no significant effect due to crew treatment, but does indicate a significant effect due to both alarm and situation treatment.

The significant alarm-by-situation interaction, however, necessitated a first level breakdown analysis by situation. The results of this analysis indicate a significant alarm effect for situation 6 only (Table 3.3.20).

The single subject data used in this analysis is the same as the data used in the update rate analysis. Since no significant difference was found between the two update rate conditions, the data was pooled. Consequently, the statistics for this treatment (single pilots without alarm) are the same as the statistics of the combined update rate statistics presents in the previous section.

In situation 6, the non-alarm two man crew subjects had a mean reaction time of 8.9 seconds and a standard deviation of 4.9 seconds. Their mean reaction time in situation 7 was 3.3 seconds, while the standard deviation 3.2 seconds.

Although Tables 3.3.19 and 3.3.20 indicate that there is no statistical difference between single pilot and crew reaction times, there is a fairly substantial difference (2.1 seconds for situation 6 and 2.2 seconds for situation 7) between the means of the data for the two crew sizes. The two man crew had the longer reaction times in both cases.

A priori it would seem reasonable to expect that with a two man crew, detection times would be decreased. The addition of the second pilot permits a more continuous ATSD monitoring capability.

Judging from cockpit (Captain to First Officer) communications, the detection times are probably just as good and possibly better than the single pilot situation. The reaction time, however, is longer because in the simulator crews have a tendency to make decisions by a committee process; i.e., they discuss the situation for a few seconds before executing the escape manuever. The fact that no statistical difference shows up can be accounted for by the fact that the data spread is fairly large.

The mean response time of the single pilots with alarm was 0.6 seconds for both situation 6 and 7. The standard deviations for these scenarios were also equal to 0.3 seconds.

The crew condition response time with alarm was 1.2 seconds for situation 6 and 2.9 seconds for situation 7. Since only one good data point was obtained for this crew and alarm treatment in situation 6, the small sample standard deviation estimate is infinite. The standard deviation of the situation 7 response times is 3.3 seconds.

A significant effect due to alarm treatment is indicated for situation 6. The average time delay from blunder commencement to initiation of escape manuever (reaction time) for alarm pilots (singles and crew) was 5.2 seconds, while the average for non-alarm pilots (singles and crew) was 11.7 seconds.

In situation 7, no significant difference was found between alarm treatments. The average reaction time for alarm pilots (singles and crews) was 11.5 seconds, while the average reaction time for non-alarm pilots was 12 seconds.

The results of the analysis of variance was used as a guide in pooling the data for the ensuing discussions (see Tables 3.4.1 and 3.4.2 for combined statistics).

In situation 6, the non-alarm subjects (singles and crews) mean reaction time is 11.5 seconds. Since the nominal impact time is 19.5 seconds, this leaves 8 seconds for the average pilot or crew to perform an escape manuever.

To encompass a one sigma distribution of pilots (i.e., 84 percent of the population), the reaction delay time increases to 14.8 seconds and leaves a 4.7 second nominal escape time.

The pilots with the emergency alarm in situation 6 had a mean reaction time of 5.2 seconds. The nominal escape time was therefore 14.3 seconds. To encompass a one sigma distribution of pilots increases the reaction time to 5.6 seconds. This leaves an escape time of 13.9 seconds.

The combined (singles and crews; alarm and non-alarm), delay time referenced to blunder commencement, for situation 7 was 11.8 seconds. This means that the average pilot had a 12.2 second nominal escape time. To encompass a one sigma distribution of pilots increases the reaction time to 14.1 seconds and decreases the nominal escape time to 9.9 seconds.

4.3 Pilot Opinion Questionnaire

The purpose of the questionnaires was to get subjective views on awareness, conflict detection, conflict resolution, and workload from the participating pilots.

In response to question number one, most pilots indicated that their overall awareness with the ATSD was excellent (43 percent) or good (39 percent). Only 18 percent thought that awareness was only fair while no pilots considered it to be poor. As can be seen in Figure 3.4.19, crew and with-alarm responses tended to be more favorable than single pilot or non-alarm responses.

In response to question two, most pilots thought that awareness with the ATSD was either much better or better (52 and 43 percent respectively) than that presently available under IMC. Only five percent of the pilots thought that ATSD-derived awareness was the same as that presently available, while no pilots thought that awareness with the display was worse or much worse than present day awareness. In this case, single pilots and non-alarm subjects tended to give more favorable responses than did crews and alarm subjects.

Most subjects thought that independent operations on closely spaced parallel runways during IMC is either acceptable (35 percent) or marginally acceptable (39 percent) when the ATSD is employed (question three). Nine percent of the pilots thought that operations under such conditions are completely safe, while fifteen percent thought that they are unacceptable. Two percent of the pilots thought that such procedures are positively dangerous. For this question, crews and alarm subjects tended to respond more favorably then did single pilots and non-alarm subjects.

Most subjects had either a high (48 percent) or a moderate (46 percent) amount of confidence in their ability to detect crossover intrusions on parallel runways with the ATSD (question four). A smaller percentage had either extremely high confidence (4 percent) or low confidence (2 percent) in their detection capabilities. Again, crew and alarm subjects tended to respond more favorably than did single pilots and non-alarm subjects.

Most pilots had a high (36 percent) or moderate (44 percent) amount of confidence in their ability to resolve parallel runway conflicts when aided by the ATSD (question five). A fair percentage were extremely confident (13 percent) while a smaller percentage (9 percent) had a low confidence. Although single pilots and non-alarm subjects had a higher percentage of "Extremely High" responses than did the crew or alarm subjects, the overall tendency was far more favorable responses from the latter two groups of pilots.

The responses to question six indicate that nearly all (93 percent) subjects thought that the inclusion of the ATSD increased workload to a higher level. Two percent thought that the workload was much higher, while five percent thought that it was the same.

CHAPTER 5

SUMMARY OF RESULTS AND CONCLUSIONS

5.1 Phase II Objectives

The work statement for the Phase II effort specifies that the primary objective is to determine if the pilot assurance value of the ATSD and the pilot's ability to detect gross system errors can be improved over the findings of Phase I by:

(1) Conducting some simulation runs with the inner loop spacing task deleted from the pilot's duties.

(2) Using a two man crew, one for inner loop tasks and one for management, communication, checklist, power and configuration changes, and conflict detection, making certain that the second pilot has a normal workload.

(3) Increasing training in the use of the ATSD and simulator over Phase I training levels to determine if additional training and familiarity improves the ability of the pilot to detect conflicts.

(4) Incorporating airborne-generated conflict alarms into the simulation and comparing the pilot's ability to detect blunders with the ability of pilots not having alarms. Prior to implementation, analyze how and what possible airborne generated algorithms might be used to alert the pilot to a potential conflict.

(5) Providing a finer final approach scale than was used in the Phase I simulations so as to provide wider spacing between closely-spaced parallel runway approach courses on the ATSD. Determine if this increases the pilot's ability to detect aircraft intrusions from the adjacent ILS.

To provide a more realistic test of the pilot's ability to detect anomalous situations, the work statement further specified that radar and imperfect navigation noises be added to the aircraft appearing on the ATSD.

All of the tasks above, except Item (5), were completed during the Phase II effort and the results are documented in this report. The scale expansion called for in Item (5) was implemented, but it caused some targets to "wrap-around". This spurious effect could not be corrected before the first subjects were tested so the expanded scale had to be eliminated from the entire test series for the sake of consistency.

The work statement also called for the development of additional scenarios for use with pilots who had already been subjects to evaluate what effect increased familiarity and added training had on their performance in detecting blunders. It was found, however, that practically every measure of performance was situation dependent. If a new set of scenarios were created for the Phase I subjects, there was no way to guarantee that the new and old scenarios were fully equivalent and that performance differences were really due to additional training and not to the scenario changes. Further, the old subjects could not be tested on the existing scenarios, because they had already been exposed to the four blunder cases in that set. As a consequence, none of the Phase I pilot subjects were used in the Phase II test series.

Two relatively minor refinements were incorporated in the Phase II display formats that were not expected to affect test results. The unit digit on ground speed was eliminated to make it compatible with the ARTS display and incidents of tag overlap that sometimes occurred on final approach were corrected.

This report has compared pilot assurance under three main experimental treatments using the definition of assurance and the measurement techniques developed for the Phase I work. The first comparison (ATC Treatment) was between the party-line communication environment of the present ATC system (Today's System) and a possible future system employing the ATSD and a discrete address communications channel. The second comparison (Spacing Task Effects) was between the performance of pilots executing an in-trail spacing task and the performance of pilots without such a task, both sets of tests using the ATSD with discrete communication. The third comparison (Crew Treatment* Effects) was between a single pilot and a two man crew, both sets of tests using the ATSD with discrete communications.

5.2 ATC Treatment Effects

This treatment compared data from Phase I party line - no ATSD tests on situations 2 and 3 (12 and 10 single pilots respectively) against the corresponding discrete address -

*See the first paragraph of Chapter 3 for a discussion of the term "treatment".

ATSD tests in Phase II employing 10 single pilots. Neither set of tests incorporated the spacing task.

The analysis of variance of information components revealed improvements in pilot awareness due to the use of the ATSD. These improvements were statistically significant for the data on aircraft position, spacing, and ground speed. A detailed breakdown by situation and aircraft sequence showed further that the improvement was only significant for specific aircraft in each scenario.

The Future ATC system had fewer null responses on the stop action quiz for the position, altitude, and ground speed information components, but Today's ATC system had fewer null responses for heading. The superiority of the Future ATC system in ground speed null responses was the only item that was statistically significant, however. The Future ATC system was also superior with respect to the landing sequence information component. The margin of superiority was statistically significant.

The ATSD proved to be much better in providing pilots with conflict detection information during the single runway blunder scenario than party-line voice. No parallel runway comparisons were made since the party-line communication channel does not allow pilots to monitor aircraft tracking performance along the adjacent ILS. Only one out of the ten subjects detected an impending conflict in situation 3 on the basis of information transmitted on the voice party-line, whereas all six no-alarm subjects with the ATSD detected the conflict prior to the point of closest approach.

One serious drawback in using radio transmissions of radar vectors as a source of conflict information is that the pilot has no way of confirming that the other aircraft are actually manuevering as directed. The only way a pilot can detect a potential conflict is if the controller generates a command that would place another aircraft in close proximity to his own craft or if the pilot of the other aircraft either fails to respond or responds incorrectly to a radar vector. Whether or not a command has been executed correctly cannot be ascertained by a pilot monitoring the party-line.

The ATSD on the other hand allows the pilot to monitor visually the actual manuevers made by other aircraft. If the pilot can effectively monitor and interpret the display, then the conflict will be detected. So far, all evidence indicates that pilots can adequately extract meaningful conflict detection information from the ATSD (see Table 3.3.24).

Pilot opinions of the ATSD were generally quite favorable. Most subjects thought that their awareness with the display was superior to that with the party-line communications channel. In addition, a surprisingly large number of pilots thought that the reduced ILS separation during IMC parallel runway operations was acceptable with an ATSD.

There confidence in being able to detect and resolve blunders under such conditions with the ATSD was generally high-to-moderate. Nearly all pilots noted an increase in workload due to the ATSD. Although not specifically asked, most pilots indicated that they thought the increase in awareness was worth the additional effort.

On the basis of these objective and subjective results, it is clear that the visual presentation of traffic information on the ATSD is more effective than the party-line aural presentation as a source of pilot awareness. It is also worth noting that professional airline pilots consistently derived more information from the party-line than non-airline pilots. For this reason, the enhancement of awareness by the ATSD would be more pronounced for non-airline pilots. In a more complex ATC environment with heavier traffic, a greater number of communications, and a more complicated route structure, the differential in performance between the two information sources would probably favor the ATSD to an even greater degree. Under such conditions, the ability to focus attention on selected aircraft and to use the ATSD as an auxiliary memory to be referenced on demand would be particularly useful.

All things considered, it is concluded that the ATSD is superior to the party-line communication channel as a source of information about other traffic and as a means for detecting conflicts, hence it would be a more than adequate replacement for the voice party-line with respect to pilot assurance.

5.3 Spacing Task Effects

This treatment compared data from the Phase I tests using discrete address communications and the ATSD, spacing task included, against the corresponding Phase II tests without the spacing task. Only single pilots were employed in the treatment. The measures derived from the stop action quiz came from runs of situations 2 (7 pilots vs. 10 pilots) and situation 3 (8 pilots vs. 10 pilots), whereas the conflict detection measures were based on runs of situations 3 (8 pilots vs. 6 pilots), 6 (7 pilots vs. 8 pilots), and 7 (7 pilots vs. 8 pilots). Subjects assisted by a computer-generated alarm, of course, could not be included in this comparison.

Virtually no difference in the estimation of information components due to the spacing task was detected. Only with respect to the heading component was there a statistically significant effect overall, the without-spacing-task subjects excelling the with-spacing-task subjects. In breaking down the data by situation and aircraft sequence, however, only the spacing component for the -1 aircraft in situation 2 shows a significant difference. The difference favors the no-spacing-task subjects. There is a statistically significant difference in the heading component for the +1 aircraft on situation 3, but the difference itself is too small to be of practical consequence.

The spacing task did not influence the estimation of landing sequence in a statistically significant way, although the without-spacing subjects generally outperformed the with-spacing subjects on this component. The percentage of null responses, however, did reflect a difference between the two spacing task treatments as well as a difference by aircraft in the way that with-task pilots responded. The without-spacing-task subjects usually had a smaller percentage of missing responses than the with-spacing task subjects for the +2 and -1 aircraft, while having a larger percentage than the with-task subjects for the +1 aircraft. This is expected since the with-task pilots are spacing themselves with respect to the +1 aircraft, hence devote more attention to it. As a consequence, the null responses on aircraft +1 by the with-spacing subjects were few while the +2 and -1 aircraft generally had a substantially larger percentage of null responses.

In breaking the null response data down by information component and aircraft sequence, it was found that the null response differences were statistically significant for the altitude component (aircraft -1) and the ground speed component (aircraft +2 and -1). The results in all three cases favor the without spacing subjects.

The data from three conflict situations (3, 6, and 7) show that the Phase I with-spacing subjects failed to detect 6 conflicts out of a total of 22 exposures, whereas the without-spacing subjects detected all 22 of the conflicts they were exposed to. This difference, however, could also be due to the fact that the without-spacing subjects received intensive training in conflict detection prior to testing and the with-spacing pilots in Phase I had no conflict experience whatsoever in their training program. With the present experimental data, therefore, it is only possible to say that the combination of more intensive conflict training and the elimination of the spacing task improves conflict detection performance.

The spacing task indirectly guaranteed that the subject's aircraft and the intruding aircraft would have a very specific kinematic relationship at the time the intrusions become acute. Without the spacing task, however, there was considerable variation in the relative position of the subject's aircraft and the intruding aircraft. As a consequence, it was not possible to use the actual miss distance during a conflict as a measure of performance and the point of closest approach was not even recorded in the tests because the data would have been meaningless. In general, subjects had no difficulty executing safe evasive maneuvers in the less acute situations 3 and 4, but in the crossover cases on the closely-spaced parallel runways (situations 6 and 7), an immediate and flawless pilot reaction is barely adequate to evade conflict regardless of the test conditions.

The inclusion of the spacing task allows a more efficient and higher capacity terminal area operation. It permits aircraft separation to be decreased under Instrument Meteorological Conditions (IMC) to approach spacings employed under Visual Meteorological Conditions (VMC) and also reduces pilot-controller communications. It does, however, detract from the pilot's ability to monitor the ATSD for information not associated with the +1 aircraft.

5.4 Crew Treatment Effects

In the evaluation of information components under the crew treatment, only the single runway data collected in Phase II was utilized. These tests employed the ATSD and a discrete address communication mode. There was no in-trail spacing task. Ten single pilots and five crews made up the subject population.

No significant difference between single pilots and two-man crews in estimating information components was observed. In all cases, however, the crews had a smaller percentage of null responses. In the breakdown of null response data by information component, the better performance of the crew was statistically significant with respect to the altitude and ground speed components. The size of the crew was significant in estimating the aircraft identification component of only the +2 aircraft of situation 4 and was not significant in estimating any of the landing sequence components.

Included in the crew treatment evaluation was a comparison of the effectiveness of alarms in aiding the pilot in blunder detection. The alarms consisted of a two mile proximity alert, employed in both single and parallel runway scenarios, and an emergency alarm based on a Tau criteria which was used on parallel runway simulations only. Tau was computed by dividing lateral distance to the adjacent ILS by lateral rate. The two mile alert indicated the presence of another aircraft within a two nautical mile radius of the subject. Penetration of this range caused a gong to ring and the target symbol to blink. The emergency audible alarm was triggered whenever an aircraft on the adjacent ILS violated a Tau threshold set at 28 seconds.

Also included in the crew treatment evaluation was a comparison of the effect of information update period (either four seconds or one second) on conflict detection capabilities. The four second period is representative of information update associated with the ARTS III system and is determined by the scan rate of the surveillance radars. The one second update period is expected to be available with the electronic scan antennas being considered for the discrete address beacon system (DABS).

The detection of conflicts prior to the point of closest approach occurred in 100% of the Phase II exposures, hence it is not possible to draw statistically significant conclusions with respect to the value of the second crew member, the alarms, or the faster update rate using only the detec n performance data. The reaction time measurements, however, did provide some useful clues.

The human reaction time in a conflict would normally consist of both controller and pilot delays, but in these experiments only the pilot was in the decision making loop. The pilot reaction time consists of three delay times; detection time, decision time, and response time. Detection time is defined as the time delay between blunder commencement and detection by the pilot. The decision delay time is defined to be the time required to decide upon an appropriate escape manuever. The pilot response time is defined to be the time it takes the pilot to initiate the chosen manuever and is caused primarily by the pilot's neuromuscular dynamics. Pilot reaction time is, therefore, the time between the commencement of the blunder and the initiation of the evasive manuever.

In the simulations conducted in Phase II, a standard escape manuever was specified. Decision time was, therefore, not a major factor. In reality, there would be a number of alternative escape routes and the pilot would have to choose among them.

For with-alarm subjects, detection time is determined by the alarm threshold. Since no decision time delay is involved, the remaining delay can be attributed to pilot response time.

In the non-alarm case, the measured delay times include both the detection and response times. An estimate of the detection times of these subjects can be made by assuming that their response times are the same as those of the alarm subjects.

In the Phase I crossover tests, the time for the intruder to fly from ILS centerline to ILS centerline was 24 seconds. This results from a continuous turn with a bank angle of 18.7° at 130 knots. Single pilots were used in these tests and no computer-generated alarm was employed. Only 4 out of 7 of the subjects detected the intrusion.

In the Phase II crossover tests run by Melanson, the 18.7° bank angle was retained and five separate test conditions were evaluated:

1. Single pilot, alarm, 4 second update (4 subjects)
2. Single pilot, no alarm, 4 second update (4 subjects)
3. Single pilot, no alarm, 1 second update (4 subjects)
4. Crew, alarm, 4 second update (2 crews)
5. Crew, no alarm, 4 second update (2 crews)

All of these subjects detected the conflict, even without the benefit of an alarm. Two factors could be responsible for the better performance shown by Melanson's pilots. First, they were given extensive training in conflict detection, whereas the Phase I subjects had no conflict experience prior to their test runs. Secondly, the spacing task was eliminated in the Phase II experiments, i.e., the Phase II pilots did not have to establish and maintain a specified spacing with respect to the aircraft ahead. This reduced their workload and allowed them to monitor the ATSD more diligently for conflicts.

The trajectories of the intruding aircraft and ownship for the Phase I and Phase II crossover with critical stagger are shown in Fig. 5.4.1. Both aircraft have a ground speed of 130 knots. The intruder has a bank angle of 18.7°

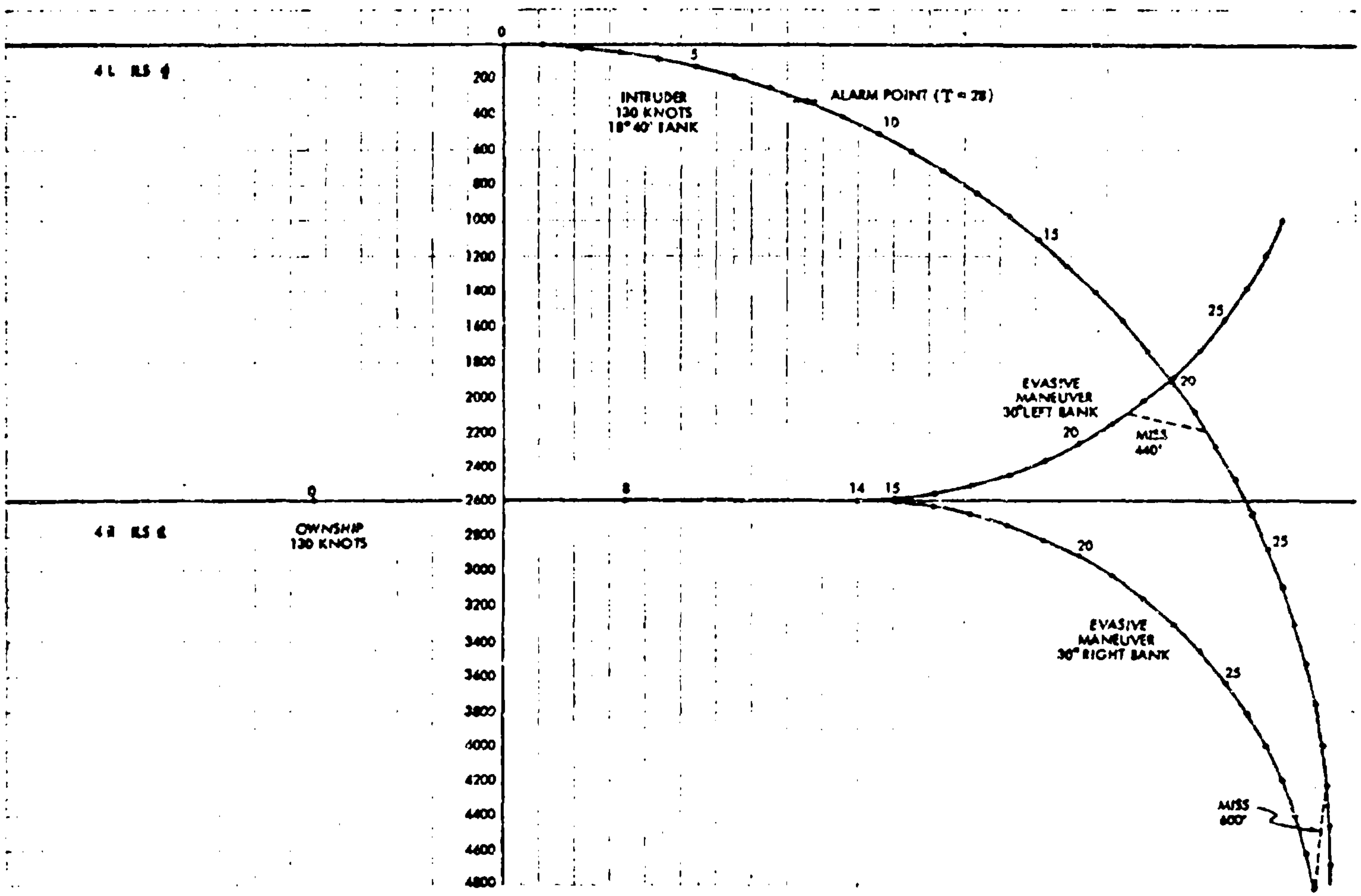


Figure 5.4.1 Typical Trajectories in Phase 2 Crossover Case

and ownship's evasive maneuver is a turn to the right with a 30° bank angle. When a crossover alarm was employed in Phase II, it went off 8 seconds after the start of the intrusion or 16 seconds before impact would occur on the adjacent ILS centerline. The pilot reaction time, defined as the time delay between the sounding of the alarm and the initiation of the subject's lateral escape maneuver, averaged 3.6 seconds for four single pilots and two crews. In addition, the 707 has a relatively slow roll rate and 5 seconds are required to establish a bank angle of 30°. Lateral offset is being developed while ownship is banking, however. The combined effects of pilot reaction time and limited roll rate are approximated in Fig. 5.4.1 by a fixed 6 second delay between the sounding of the alarm and the initiation of the circular escape trajectory. The miss distance for the flight path shown is 600 feet. Had any of the subjects turned left instead of right, the miss distance would have been only 440 feet. Thus, the pilots in the Phase II crossover test had a reasonable chance of developing a "safe" miss distance even when the intrusion occurred at the critical point, i.e., the stagger at which a collision would take place on the ILS centerline if no evasive manuever were attempted. Since the spacing task was eliminated in the Phase II tests the dispersion of ownship position about the critical point at the start of the intrusion was great and several pilots achieved even greater miss distances than that illustrated in Fig. 5.4.1. As a consequence, many subjects on the opinion questionnaire rated independent operation with closely-spaced runways as acceptable using the ATSD. For example, 60% of the pilots had "high" confidence in being able to detect potential conflicts during closely-spaced parallel IFR operations using the ATSD with an alarm. Fifty percent had "high" confidence that they could resolve conflicts successfully under these conditions. Seventy percent rated the safety associated with closely-spaced IFR operations using the ATSD with an alarm as either completely safe or acceptable.

These results are misleading, however, because the crossover maneuver employed on the Phase II subjects was by no means a worst case. The intruding aircraft had a bank angle of only 18.7°, whereas bank angles in excess of 45° would be possible without stalling. The crossover time decreases with increased speed, hence a higher final-approach speed than 130 knots would represent a more realistic worst case. Finally, the alarm threshold in the Phase II tests corresponds to an τ of 28, τ being defined as the intruding aircraft's lateral distance to the adjacent localizer divided by its lateral rate. Normal ILS tracking would frequently generate lower values of τ than this, hence a τ of 28 must be regarded as impractical because it would cause at high false alarm rate.

Consequently, at the end of the Phase II program, it was felt that a definitive answer to the question, "Under what conditions would independent operations on close-spaced parallel runways be acceptable to pilots?" was still lacking. Granted that a crossover would be a rare event, it was felt that pilots would accept closely-spaced parallels if they had a reasonable chance of detecting and evading a true worst case intrusion.

Pilot reaction times were measured under all the crew, alarm and update rate treatments in parallel runway situations 6 and 7 (Table 3.2.6). The statistical results indicate that the effect on reaction time due to update rate was not significant, but that the emergency alarm had a significant effect with both single pilots and crews on the reaction time for situation 6. Reaction times were found to be situation dependent for both update rate and alarm treatments.

The emergency alarm appreciably cut down the pilot reaction time in situation 6 (ILS acquisition overshoot), but offered no improvement for situation 7 (ILS crossover). The conclusion to be drawn from this is that the pilot's ability to monitor the ATSD coupled with his own internal detection threshold was equivalent in performance on the average to the automatic Tau alarm for situation 7, but was inferior to the automatic alarm for situation 6. The blunder in situation 7 is easier for the pilot to detect because of the abrupt course change.

The equivalence noted in situation 7 is, of course, true only for the Tau threshold (28) employed in these tests. The selection of a higher threshold would probably cause a statistical difference in this case as well.

No statistical difference in reaction time due to crew treatment was detected for either situation. The two man crews did, however, tend to react slower than did the single pilots. This was caused by a tendency for crews to discuss the situation briefly before acting. This, therefore, is an area where more work needs to be done in refining procedures and training crews.

Although clear, statistically significant differences between the performance of crews and single pilots were lacking, it is worth noting that the crews generally rated the ATSD higher on the opinion questionnaire than the single pilots. This probably reflects the lower workload that the ATSD imposed on the crew since the inner-loop and outer-loop tasks were split between the two pilots instead of being handled by one man.

5.5. Final Comment

With the contemplated implementation of discrete address communication channels and the initiation of independent operations on parallel runways spaced closer than 5000 feet, it would appear that consideration should be given to developing a means for providing pilots with assurance (peace of mind) that ATC operations are running smoothly and safely. With present terminal area surveillance and monitoring procedures, independent operations under IMC are conducted with parallel runways separations down to 5,000 feet. It appears that improved surveillance will permit separations down to 3,500 feet while the desired ultimate objective is 2,500 feet. Providing a pilot with a means of cross-checking the operations of the ATC system and the performance of other pilots appears to be one method of enhancing safety and generating pilot acceptance of closer parallel runway operations. The ATSD is one practical mechanism for providing the pilot with an easy-to-use real-time picture of the ATC situation around him.

APPENDIX*

STATISTICAL ANALYSIS PROCEDURES EMPLOYED IN STUDY

A.1 Analysis of Variance

The arithmetic mean or average of a set of N measurements having the values X_i is given by

$$M = \frac{\sum_{i=1}^{N} X_i}{N} \tag{1}$$

The mean square of the data points about this average value is called the variance of the distribution. It is given by

$$s^2 = \frac{\sum_{i=1}^{N} (X_i - M)^2}{N} \tag{2}$$

The operation of summing squares about a mean value occurs repeatedly in the analysis of variance. A more convenient formula for performing this operation is the following:

$$\sum_{i=1}^{N} (X_i - M)^2 = \sum_{i=1}^{N} X_i^2 - \frac{(\sum_{i=1}^{N} X_i)^2}{N} \tag{3}$$

from which $$s^2 = \frac{\sum_{i=1}^{N} X_i^2}{N} - \left(\frac{\sum_{i=1}^{N} X_i}{N}\right)^2 = \frac{\sum_{i=1}^{N} X_i^2}{N} - M^2 \tag{4}$$

*Contributed by Mark E. Connelly

In general, a set of data points is a sample from a very much larger set of data points which represent the entire population being evaluated. The use of samples introduces random errors which could be eliminated if we had the patience and resources to test the entire population. This is seldom possible in practice. The techniques of statistical analysis have been developed to enable us to make judgements with reasonable assurance on the basis of data contaminated by errors due to a finite sample, the presence of experimental factors other than those we are trying to evaluate, measurement errors, variations in the subjects or experimental material, and variations in the ambient conditions under which the experiment is conducted. Specifically, the analysis of variance technique breaks up the total sum of squares about the grand average of all data points into two or more component parts. In the simplest case, there are only two components, a between-groups sum of squares, which measures the variation of the group means around the over-all mean, and a within-groups sum of squares, which measures the variation of the data within each of the groups around the respective group means. Each of these component sums of squares is divided by an appropriate degree of freedom index to yield two independent estimates of the population variance. The ratio of these two estimates is called the F-ratio. If there is no significant difference between the two groups of data, the value of F, on the average, will be 1.00. However, if the group means differ significantly from each other because of the effect of the experimental variable, the between-groups variance will be greater than the within-groups variance and F will be considerably larger than 1.00. Standard tables of the F distribution enable us to determine the probability that a given value of F could have occurred by chance. In this study, an experimental treatment is considered statistically significant if the corresponding F-ratio has less than a 5% probability of occurring by chance.

To illustrate the technique, a step-by-step explanation of the three-way analysis of variance employed in this report will be presented. Beyond that, a more detailed discussion of the theory and practice of analysis of variance may be found in References 8, 10, 11, 12, 13 and 14.

As our illustrative example, we will repeat the three way analysis of variance shown in Table 3.3.3 and find F-ratios for the spacing error information component. The main effects are crew, situation, and aircraft sequence. The data for the analysis is derived from Table 3.2.2 and is restated in Table A.1 in a form more convenient for computation. In generating Table A.1 from Table 3.2.2, Equation (1) is used to find $\sum X_i$ values and Equation (4) is used to find the $\sum X_i^2$ values. The number of samples in each category (N) is obtained by subtracting the number of

TABLE A.1

DATA FOR THREE-WAY ANALYSIS OF VARIANCE
(SPACING ERROR INFORMATION COMPONENT)

	Single Pilot											
	Situation #1			Situation #2			Situation #3			Situation #4		
Aircraft Sequence	+2	+1	-1	+2	+1	-1	+2	+1	-1	+2	+1	-1
N	8	10	9	9	9	7	9	9	9	10	10	10
$\sum_{i=1}^{N} x_i$	12	14	-23.4	6.3	-3.6	9.1	-6.3	-2.7	-24.3	8	14	40
$\sum_{i=1}^{N} x_i^2$	43.9	29.6	81.1	36.9	62.3	39.6	65.3	26.8	131.2	38.8	46.5	268.9

	Crew											
	Situation #1			Situation #2			Situation #3			Situation #4		
Aircraft Sequence	+2	+1	-1	+2	+1	-1	+2	+1	-1	+2	+1	-1
N	5	5	5	4	4	4	5	5	5	5	5	4
$\sum_{i=1}^{N} x_i$	0	5	-12.5	-1.6	2.8	5.2	-8	-5	-14	.5	6	19.6
$\sum_{i=1}^{N} x_i^2$	1.8	7.5	41.1	34.3	21.3	38.1	14.6	7.5	84.2	16.3	18.5	173.5

null responses from the number of subjects. The individual subjects contributing to each category have been pooled in Table A.1

A <u>between-groups sum of squares</u> is obtained by replacing each item by its own group average and taking the sum of squares about the grand average of all groups. For example, suppose we have three groups of data X, Y, and Z consisting of N_1, N_2, and N_3 samples respectively.

$$\text{X average} = \frac{\sum_{i=1}^{N_1} X_i}{N_1} = M_1 \tag{5}$$

$$\text{Y average} = \frac{\sum_{i=1}^{N_2} Y_i}{N_2} = M_2 \tag{6}$$

$$\text{Z average} = \frac{\sum_{i=1}^{N_3} Z_i}{N_3} = M_3 \tag{7}$$

$$\text{grand average} = \frac{\sum_{i=1}^{N_1} X_i + \sum_{i=1}^{N_2} Y_i + \sum_{i=1}^{N_3} Z_i}{N_1 + N_2 + N_3} = M_T \tag{8}$$

Replacing each item by its own group average and taking the sum of squares about the grand average, we have

Between Group Sum of Squares

$$= N_1(M_1 - M_T)^2 + N_2(M_2 - M_T)^2 + N_3(M_3 - M_T)^2$$

$$= N_1(M_1^2 - 2M_1M_T + M_T^2) + N_2(M_2^2 - 2M_2M_T + M_T^2) + N_3(M_3^2 - 2M_3M_T + M_T^2)$$

$$= (N_1M_1^2 + N_2M_2^2 + N_3M_3^2) - 2M_T(N_1M_1 + N_2M_2 + N_3M_3) + M_T^2(N_1 + N_2 + N_3) \quad (9)$$

but from Equations 5, 6, 7 and 8

$$N_1M_1 + N_2M_2 + N_3M_3 = \sum_{i=1}^{N_1} X_i + \sum_{i=1}^{N_2} Y_i + \sum_{i=1}^{N_3} Z_i = M_T(N_1 + N_2 + N_3) \quad (10)$$

hence

Between Group Sum of Squares

$$= (N_1M_1^2 + N_2M_2^2 + N_3M_3^2) - M_T^2(N_1 + N_2 + N_3)$$

$$= \frac{\left(\sum_{i=1}^{N_1} X_i\right)^2}{N_1} + \frac{\left(\sum_{i=1}^{N_2} Y_i\right)^2}{N_2} + \frac{\left(\sum_{i=1}^{N_3} Z_i\right)^2}{N_3} - \frac{\left(\sum_{i=1}^{N_1} X_i + \sum_{i=1}^{N_2} Y_i + \sum_{i=1}^{N_3} Z_i\right)^2}{N_1 + N_2 + N_3} \quad (11)$$

Equation (11) applies strictly to the three data group case, but it illustrates the standard form for computing the between-group sum of squares. In the general case, the sum of all items in each group is squared and divided by the number of samples in that group. These values are added together and the correction factor subtracted from the sum. The correction factor, which occurs repeatedly in an analysis, is the square of the sum of all items in all groups divided by the total number of samples. In our illustrative analysis, the correction factor is given by

$$\text{correction factor} = \frac{(41.1)^2}{165} = 10.24$$

The F-ratios for the main effects and for first order and second order interactions are obtained from six tables derived from Table A.1 as follows:

Table A.2

(Obtained from A.1 by summing $\sum X_i$ values over Aircraft)

	Single Pilot	Crew
Situation #1	2.6 27 samples	-7.5 15 samples
Situation #2	11.8 25 samples	6.4 12 samples
Situation #3	-33.3 27 samples	-27 15 samples
Situation #4	62 30 samples	26.1 14 samples

Table A.3

(Obtained from A.1 by summing $\sum X_i$ values over Situations)

	Single Pilot	Crew
Aircraft +2	20 36 samples	-9.1 19 samples
Aircraft +1	21.7 38 samples	8.8 19 samples
Aircraft -1	1.4 35 samples	-1.7 19 samples

Table A.4

(Obtained from A.1 by summing $\sum X_1$ values over crew sizes)

	Situation #1	Situation #2	Situation #3	Situation #4
Aircraft +2	12 13 samples	4.7 13 samples	-14.3 14 samples	8.5 15 samples
Aircraft +1	19 15 samples	-.8 13 samples	- 7.7 14 samples	20 15 samples
Aircraft -1	-35.9 14 samples	14.3 11 samples	-38.3 14 samples	59.6 14 samples

Table A.5

(Obtained from A.2 by summing over situations)

Single Pilot	Crew
43.1 109 samples	-2.0 56 samples

Grand Total 41.1
165 samples

Table A.6

(Obtained from A.3 by summing over crew sizes)

Aircraft +2	Aircraft +1	Aircraft -1
10.9 55 samples	30.5 57 samples	-0.3 53 samples

Grand Total 41.1
165 samples

Table A.7

(Obtained from A.4 by summing over aircraft)

Situation #1	Situation #2	Situation #3	Situation #4
-4.9 42 samples	18.2 37 samples	50.3 42 samples	88.1 44 samples

Grand Total 41.1
165 samples

With this background and the data presented in Tables A.1 through A.7, we can now calculate the sum of squares and degrees of freedom associated with the crew effect(C), the aircraft effect (A), and the situation effect (S), as well as the C x A, C x S, A x S, and C x A x S interactions.

Crew Effect

Divide the square of each entry in Table A.5 by the number of samples contributing to the entry, and subtract the correction factor from the sun of these terms.

$$\text{Crew Effect Sum of Squares} = \frac{(43.1)^2}{109} + \frac{(-2)^2}{56} - 10.24 = 6.87$$

Since there are two levels of the crew factor, there is one degree of freedom associated with this sum of squares. The corresponding variance estimate is the sum of squares divided by the degrees of freedom.

$$\text{Variance Estimate} = \frac{6.87}{1} = 6.87$$

Aircraft Effect

Divide the square of each entry in Table A.6 by the number of samples contributing to the entry, and subtract the correction factor from the sum of these terms.

$$\text{Aircraft Effect Sum of Squares} = \frac{(10.9)^2}{55} + \frac{(30.5)^2}{57} + \frac{(-0.3)^2}{53} - 10.24 = 8.24$$

There are three levels of the aircraft factor, hence there are two degrees of freedom.

$$\text{Variance Estimate} = \frac{8.24}{2} = 4.12$$

Situation Effect

Divide the square of each entry in Table A.7 by the number of samples contributing to the entry, and subtract the correction factor from the sum of these terms.

Situation Sum of Squares

$$= \frac{(-4.9)^2}{42} + \frac{(18.2)^2}{37} + \frac{-60.3)^2}{42} + \frac{(88.1)^2}{44} - 10.24 = 262.26$$

There are three degrees of freedom corresponding to the four levels of the situation factor.

$$\text{Variance Estimate} = \frac{262.26}{3} = 87.4$$

Crew X Aircraft Interaction

Divide the squares of each entry in Table A.3 by the number of samples contributing to the entry. From the sum of these terms subtract the sum of squares calculated for the crew effect, the sum of squares calculated for the aircraft effect, and the correction factor.

C x A Sum of Squares

$$= \frac{(20)^2}{36} + \frac{(21.7)^2}{38} + \frac{(1.4)^2}{35} + \frac{(-9.1)^2}{19} + \frac{(8.8)^2}{19} + \frac{(-1.7)^2}{18}$$

$$- 6.87 - 8.24 - 10.24 = 6.8$$

The crew effect had one degree of freedom and the aircraft effect had two degrees of freedom, therefore the C x A interaction has 1 x 2 = 2 degrees of freedom

$$\text{Variance Estimate} = \frac{6.8}{2} = 3.4$$

Crew x Situation Interaction

Divide the square of each entry in Table A.2 by the number of samples contributing to the entry. From the sum of these terms subtract the sum of squares calculated for the crew effect, the sum of squares calculated for the situation effect, and the correction factor.

C x S Sum of Squares

$$= \frac{(2.6)^2}{27} + \frac{(11.8)^2}{25} + \frac{(-33.3)^2}{27} + \frac{(62)^2}{30} + \frac{(-7.5)^2}{15} + \frac{(6.4)^2}{12}$$

$$+ \frac{(-27)^2}{15} + \frac{(26.1)^2}{14} - 6.87 - 262.26 - 10.24 = .07$$

The crew effect had one degree of freedom and the situation effect had three degrees of freedom, hence the C x S interaction has 1 x 3 = 3 degrees of freedom.

$$\text{Variance Estimate} = \frac{.07}{3} = .02$$

Aircraft x Situation Interaction

Divide the square of each entry in Table A.4 by the number of samples contributing to the entry. From the sum of these terms, subtract the sum of squares calculated for the aircraft effect, the sum of squares calculated for the situation effect, and the correction factor.

A x S Sum of Squares

$$= \frac{(12)^2}{13} + \frac{(19)^2}{15} + \frac{(-35.9)^2}{14} + \frac{(4.7)^2}{13} + \frac{(-.8)^2}{13} + \frac{(14.3)^2}{11}$$

$$+ \frac{(-7.7)^2}{14} + \frac{(-38.3)^2}{14} + \frac{(8.5)^2}{15} + \frac{(20)^2}{15} + \frac{(59.6)^2}{14}$$

$$- 8.24 - 262.26 - 10.24 = 275.6$$

The aircraft effect had two degrees of freedom and the situation effect had three degrees of freedom, therefore the A x S interaction has 2 x 3 = 6 degrees of freedom.

$$\text{Variance Estimate} = \frac{275.6}{6} = 45.9$$

Crew x Aircraft x Situation Interaction

Divide the square of each $\sum X_i$ entry in Table A.1 by the number of samples contributing to the entry (N). From the sum of these terms, subtract the sums of squares calculated for the crew effect, the aircraft effect, the situation effect, the C x A interaction, the C x S interaction, the A x S interaction, and the correction factor.

C x A x S Sum of Squares

$$= \frac{(12)^2}{8} + \frac{(14)^2}{10} + \ldots + \frac{(19.6)^2}{4} - 6.87 - 8.24 - 262.26$$

$$- 6.8 - .07 - 275.6 - 10.24 = 8.77$$

The degrees of freedom are 1 x 2 x 3 = 6

$$\text{Variance Estimate} = \frac{8.77}{6} = 1.46$$

Total Sum of Squares and Residuals

The total sum of squares of all data points about the grand mean, from Equation (3), is given by the sum of all $\sum X_i^2$ entries in Table A.1 minus the correction factor.

Total Sum of Squares

$$= 43.9 + 29.6 + 81.1 + \ldots + 173.5 - 10.24 = 1321.6$$

Since Table A.1 represents 165 data points, the total degrees of freedom will be 164. To get the residual sum of squares, we substract from the total sum of squares the sums of squares calculated for crew, aircraft, situation, C x A interaction, C x S interaction, A x S interaction, and the C x A x S interaction.

Residual Sum of Squares

$$= 1321.6 - 6.87 - 8.24 - 262.26 - 6.8 - .07 - 275.6 - 8.77 = 753$$

The residual degrees of freedom are found by subtracting the sum of the degrees of freedom of all the main effects and the interactions from the total degrees of freedom.

Residual Degrees of Freedom

$= 164 - 1 - 2 - 3 - 2 - 3 - 6 - 6 = 141$

$$\text{Residual Variance Estimate} = \frac{753}{141} = 5.34$$

On the basis of the calculations made thus far, we can form the final Table of Analysis of Variance.

Table A.8

	Effect	Sum of Squares	Degrees of Freedom	Variance Estimate	F-Ratio	Significant at 5% Level
Main Factors	Crew	6.87	1	6.87	1.33	No
	Aircraft	8.24	2	4.12	—	—
	Situation	262.26	3	87.4	—	—
Inter-actions	A x S	275.6	6	45.9	8.86	Yes
	C x S	.07	3	.02	.00	No
	C x A	6.8	2	3.4	.66	No
	C x A x S	8.77	6	1.46	.27	No
Experi-mental Error	Residual	753	141	5.34		

The statistical significance of the F-ratio for the C x A x S interaction must be tested first. If this second order interaction is significant, it is not valid to test any of the first order interactions (C x A, C x S, or A x S) against the residual.

The F-ratio for the C x A x S interaction is calculated by dividing the C x A x S variance estimate by the residual variance estimate.

$$\text{F-Ratio for C x A x S Interaction} = \frac{1.46}{5.34} = .27$$

Consulting a standard F-distribution table and entering the table at the appropriate degrees of freedom for the numerator and denominator (6:141), we find that an F value of at least 2.18 is needed to reach the 5% level, i.e., the sum of squares had more than a 5% probability of occurring by chance. The C x A x S interaction, therefore, is not statistically significant.

Since the C x A x S interaction is not significant, we can lump the C x A x S sum of squares and degrees of freedom with the corresponding residual values to get a slightly more accurate estimate of the residual variance. From Table A.8, we have

Revised Residual Sum of Squares = 753 + 8.77 = 761.8

Revised Residual Degrees of Freedom = 141 + 6 = 147

$$\text{Revised Residual Variance Estimate} = \frac{761.8}{147} = 5.18$$

The revised variance is used in the calculation of F-ratios for the first order interactions C x A, C x S, and A x S. Each F-ratio is obtained by dividing the corresponding interaction variance estimate by the residual variance estimate.

The results are indicated in Table A.8. As one might expect from the non-parallel plots in Fig. 3.4.13, the A x S interaction is highly significant, hence it is not valid to test either the aircraft or situation main effects against the residual. To evaluate these effects, it is necessary to break down the analysis of variance into four separate analyses for each of the four situations. The results of this breakdown are given in Table 3.3.5. They show that aircraft sequence is significant in estimating spacing in situations 1, 3, and 4, but that crew size is not significant in any of the situations.

Since neither the C x A or C x S interaction was significant in Table A.8, the crew main effect can be tested against the revised residual. The resultant F-ratio (1.33) is not significant. The numbers presented in Table A.8 differ slightly from those in Table 3.3.3 because the latter were derived from raw experimental data whereas the former were derived from the means and standard deviations given in Table 3.2.2. Note that these small differences did not affect the final conclusions since the analysis of variance is relatively insensitive to minor perturbations in the input data.

A.2 Chi-Square Tests

The X^2 (chi-square) test tells us whether the observed frequency of some event differs significantly from the frequency which might be expected according to some assumed hypothesis. Corresponding to each frequency predicted by the hypothesis (E for expected), there will be an experimentally observed frequency (O for observed). If there are N items in the contingency table, the index X^2 is calculated by the following sum of terms.

$$X^2 = \sum_{i=1}^{N} \frac{(O_i - E_i)^2}{E_i} \qquad (12)$$

Naturally, if the observed frequency equals the expected frequency in every case, the value of X^2 is zero and the hypothesis is upheld. For finite values of X^2, tables of the X^2 distribution give us the probability that the observed frequencies occurred by chance even though the hypothesis is actually true. To use these tables, one must also calculate the degrees of freedom associated with the X^2 value. The degrees of freedom are equal to the number of entries in the contingency table whose frequency may be assigned arbitrarily. If the contingency table has r rows and p columns, the degrees of freedom are (r-1) (p-1).

We will illustrate the X^2 test by using it to determine if the crew treatment had a significant effect on null responses. A null response is a missing data point, i.e., an estimate of aircraft position, altitude, ground speed, or heading that a subject failed to record on his stop-action quiz. The raw data for this test (Table A.9) was derived

TABLE A.9

NULL RESPONSES (CREW TREATMENT)

	Single Pilot												Crew												
	Situation #1			Situation #2			Situation #3			Situation #4			Situation #1			Situation #2			Situation #3			Situation #4			
Aircraft Sequence →	+2	-1	-1	+2	+1	-1	+2	+1	-1	+2	+1	-1	+2	+1	-1	+2	+1	-1	+2	+1	-1	+2	+1	-1	Totals
Position Responses	8	10	9	10	10	8	9	9	9	10	10	10	5	5	5	4	4	4	5	5	5	5	5	4	168
Position Nulls	2	0	1	0	0	2	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	9
Altitude Responses	5	9	7	9	8	9	6	9	8	10	10	5	4	5	5	4	4	3	5	5	5	5	5	5	151
Altitude Nulls	5	1	3	1	2	1	4	1	2	0	0	4	1	0	0	0	0	1	0	0	0	0	0	0	26
Ground Speed Responses	5	8	7	8	9	9	8	10	7	10	9	6	4	5	5	4	4	4	5	5	4	5	5	5	151
Ground Speed	5	2	3	2	1	1	2	0	3	0	1	4	1	0	0	0	0	0	0	0	1	0	0	0	26
Heading Responses	6	7	7	6	6	6	7	7	6	7	7	6	3	4	4	2	3	3	4	4	4	3	4	3	119
Heading Nulls	4	3	3	4	4	4	3	3	4	[illegible]	3	4	2	1	1	2	1	1	1	1	1	2	1	2	58
Total Responses	24	34	30	33	33	32	30	35	30	37	36	28	16	19	19	14	15	14	19	19	18	18	19	17	589
Total Nulls	16	6	10	7	7	8	10	5	10	3	4	12	4	1	1	2	1	2	1	1	2	2	1	3	119
Total	40	40	40	40	40	40	40	40	40	40	40	40	20	20	20	16	16	16	20	20	20	20	20	20	708

from the null response data in Tables 3.2.1, 3.2.3, 3.2.4, and 3.2.5.

We first test the hypothesis that there is no difference in the frequency of null responses due to situations (all aircraft pooled). This is really eight separate X^2 tests based on data for the four information components and the two crew sizes. The observed and expected null frequencies and response frequencies are given in Table A.10. The method of calculating expected values can be explained in terms of the single pilot position null entries.

Since there are a total of 8 position nulls with single pilots, one would expect, in view of the situation hypothesis, that each of the four situations would have 2 position nulls. A total of 30 position responses are possible in each of the four situations, so if the expected number of null responses is 2, the expected number of responses must be 28.

Similarly, the 24 altitude nulls yield an expected value of 6 nulls for each of the four situations and an expected 24 altitude responses. This relatively simple method of calculating the expected number of nulls and responses cannot be used on the crew data, however, because 4 crews were tested in Situation 2 and 5 crews were tested in Situations 1, 3, and 4. For each information component, a total of 57 replies was possible, 15 from Situation 1, 12 from Situation 2, 15 from Situation 3, and 15 from Situation 4. Since the crews produced only one position null overall, the expected number of position nulls (E) in each situation is given by:

Situations #1,3,4 $E = \frac{1}{57} 15 = .26$

Situation #2 $E = \frac{1}{57} 12 = .21$

A total number of 15 position responses was possible in Situations 1, 3, and 4, hence the expected number of responses is 15 minus .26 or 14.74. Only 12 position responses were possible in Situation 2, so the expected number of responses is 12 minus .21 or 11.79. The expected number of nulls and responses for the other crew information components is computed in a similar fashion and the results are given in Table A.10.

Using the observed and expected frequencies in Table A.10, we now compute eight separate X^2 values corresponding to the eight combinations of crew size and information components. To illustrate the procedure, the X^2 value for single-pilot position null data is obtained from Eq. 12 as follows:

$$X^2(\text{single pilot, position nulls}) = \frac{(3-2)^2}{2} + \frac{(3-2)^2}{2} + \frac{(0-2)^2}{2} + \frac{(27-28)^2}{28} + \frac{(27-28)^2}{28} + \frac{(30-28)^2}{28} = 3.21$$

TABLE A.10

NULL RESPONSE DATA TO TEST SITUATION HYPOTHESIS
(AIRCRAFT POOLED)

Crew Size	Information Component		Situation #1	Situation #2	Situation #3	Situation #4
Single Pilot	Position Nulls	O	3	2	3	0
		E	2	2	2	2
	Position Responses	O	27	28	27	30
		E	28	28	28	28
	Altitude Nulls	O	9	4	7	4
		E	6	6	6	6
	Altitude Responses	O	21	26	23	26
		E	24	24	24	24
	Ground Speed Nulls	O	10	4	5	5
		E	6	6	6	6
	Ground Speed Responses	O	20	26	25	25
		E	24	24	24	24
	Heading Nulls	O	10	12	10	10
		E	10.5	10.5	10.5	10.5
	Heading Responses	O	20	19	20	20
		E	19.5	19.5	19.5	19.5
Crew	Position Nulls	O	0	0	0	1
		E	.26	.21	.26	.26
	Position Responses	O	15	12	15	14
		E	14.74	11.79	14.74	14.74
	Altitude Nulls	O	1	1	0	0
		E	.53	.42	.53	.53
	Altitude Responses	O	14	11	15	15
		E	14.47	11.58	14.47	14.47
	Ground Speed Nulls	O	1	0	1	0
		E	.53	.42	.53	.53
	Ground Speed Responses	O	14	12	14	15
		E	14.47	11.58	14.47	14.47
	Heading Nulls	O	4	4	3	5
		E	4.21	3.37	4.21	4.21
	Heading Responses	O	11	8	12	10
		E	10.79	8.63	10.79	10.79
Total Nulls		O	38	27	29	25
		E	30.25	28.24	30.25	30.25
Total Responses		O	142	141	151	155
		E	149.75	139.76	149.75	149.75

Entering the X^2 table for 3 degrees of freedom, we find that the distribution of position replies had a 25-50% probability of occurring by chance, i.e., the situation hypothesis for this null component is upheld. The situation hypothesis is likewise upheld for the other seven information components as shown in Table 3.3.16 (a).

In actual practice, the X^2 test cannot be safely applied if the expected frequency in any cell is less than 5. Because of this, all the data in Table A.10 should be pooled and the X^2 test applied to the totals in each column as follows.

$$X^2 = \frac{(38-30.25)^2}{30.25} + \frac{(27-28.24)^2}{28.24} + \frac{(29-30.25)^2}{30.25} + \frac{(25-30.25)^2}{30.25}$$

$$+ \frac{(142-149.75)^2}{149.75} + \frac{(141-139.76)^2}{139.76} + \frac{(151-149.75)^2}{149.75}$$

$$+ \frac{(155-149.75)^2}{149.75} = 3.61$$

With 3 degrees of freedom, the probability that the null response distribution was due to chance is between 25 and 50%, so the situation hypothesis is upheld by the pooled data also.

In the same fashion, the hypothesis that there are no differences in null responses due to aircraft sequence is upheld as shown in Table 3.3.16 (b). Since both the situation and aircraft hypotheses appear to be true, the data from all situations and all aircraft can be pooled in testing the hypothesis of most interest, namely, that the number of null responses is not affected by crew size. Table A.11 presents the data needed for this test.

In Table A.11, Yates' correction has been applied to the observed values. This is an empirical adjustment that compensates for the fact that the continuous X^2 distribution is being used for a binomial type of problem which is, in fact, discontinuous. The Yates correction decreases by .5 those observed values in the table which exceed the expected value and increase by .5 those observed values which are less than the expected value. From Eq. 12, the X^2 value for each information component can be computed individually or all the information components can be pooled and a X^2 index computed from the totals. To illustrate the general procedure once again, the X^2 calculation for the position null response data is presented below in detail:

$$X^2(\text{position nulls}) = \frac{(7.5-6.10)^2}{6.10} + \frac{(1.5-2.9)^2}{2.9} + \frac{(112.5-113.9)^2}{113.9} + \frac{(55.5-54.1)^2}{54.1} = 1.05$$

Degrees of Freedom = 1

$.25 < P < .50$ Crew hypothesis is upheld by position data

Table A.11

Null Response Data to Test Crew Hyrothesis
(Situations and Aircraft Pooled)

Information Components		Single Pilot	Crew
Position Nulls	O E	8 - .5 = 7.5 6.10	1 + .5 = 1.5 2.90
Position Responses	O E	112 + .5 = 112.5 113.9	56 - .5 = 55.5 54.1
Altitude Nulls	O E	24 - .5 = 23.5 17.63	2 + .5 = 2.5 8.37
Altitude Responses	O E	96 + .5 = 96.5 102.37	55 - .5 = 54.5 48.63
Ground Speed Nulls	O E	24 - .5 = 23.5 17.63	2 + .5 = 2.5 8.37
Ground Speed Responses	O E	96 + .5 = 96.5 102.37	55 - .5 = 54.5 48.63
Heading Nulls	O E	42 - .5 = 41.5 39.32	16 + .5 = 16.5 18.68
Heading Responses	O E	78 + .5 = 78.5 80.68	41 - .5 = 40.5 38.32
Total Nulls	O E	98 - .5 = 97.5 80.68	21 + .5 = 21.5 38.32
Total Responses	O E	382 + .5 = 382.5 399.32	207 - .5 = 206.5 189.68

In like manner, we can calculate X^2 for the other information components to obtain:

$$X^2 \text{ (altitude and ground speed nulls)} = 7.12$$

$.005 < P < .01$ Crew hypothesis rejected by altitude and ground speed data

$$X^2 \text{ (heading nulls)} = .56$$

$.25 < P < .50$ Crew hypothesis upheld by heading data

We can also pool all the information components and calculate X^2 from the totals in each column of Table A.11 with the results below:

$$X^2 \text{ (pooled data)} = 13.09$$

$$P < .005$$

crew hypothesis rejected by pooled information component data

The rejection of the crew hypothesis, of course, imples that crew size does have a significant effect on null responses in the stop action quiz. In this case, the two-man crew has fewer null responses than the single pilot and the difference is statistically significant.

REFERENCES

1. Radio Corporation of America, TELERAN, Camden, New Jersey, September 1, 1946

2. Sluka, A.L., Dynamic Simulation Studies of Pictorial Navigation Displays as Aids to Air Traffic Control in Low-Density Terminal Area and in an En-Route Area, NAFEC Project FAA-115-703X, AD-608 952, Atlantic City, N.J., February, 1963

3. Feasibility Test of Televised Radar, NAFEC Project 241-023-01X, RD-66-6 Atlantic City, N.J., January, 1966.

4. Bush, R.W., Blatt, H., Brady, F.X., A Cockpit Display of Selected NAS/ARTS Data, Lincoln Laboratory Technical Note 1970-39, Lexington, Massachusetts, December, 1970

5. Imrich, Thomas, Concept Development and Evaluation of Airborne Traffic Displays, Flight Transportation Laboratory Report R71-2, Massachusetts Institute of Technology, June, 1971.

6. Anderson, R.E., Format Evaluation for An Airborne Air Traffic Situation Display, Report MVT-71-2; M.I.T., June, 1971.

7. Howell, J.D., Simulator Evaluation of Pilot Assurance Derived from an Airborne Traffic Situation Display, FAA Report No. FAA-EM-72-3, February 1972.

8. Moroney, M.J., Facts From Figures, Penguin Books Inc., Baltimore, Md., 1968, pp. 249-269.

9. Moroney, pp. 371-457.

10. Cochran and Cox, Experimental Designs, 2nd Edition, John Wiley & Sons, New York, 1957

11. Fisher, R.A., Statistical Methods for Research Workers, 10th Edition, Oliver & Boyd, Edinburgh, 1946

12. Fisher, R.A., The Design of Experiments, 4th Edition, Oliver & Boyd, Edinburgh, 1947

13. Snedecor, G.W., Statistical Methods, 4th Edition, Iowa State College Press; Ames, Iowa; 1946

14. Kendall, M.G., The Advanced Theory of Statistics, Vol. 2, Charles Griffin & Co., London, 1946.