Spring 1988

Investigation of Display Issues Relevant to the Presentation of Aircraft Fault Information

Donald Mark Allen
Old Dominion University

Investigation of Display Issues Relevant to the Presentation of Aircraft Fault Information

by

Donald Mark Allen

B.A. May, 1983, University of Maryland, Baltimore County

M.A. May, 1985, The College of William and Mary

A Dissertation Submitted to the Faculty of Old Dominion University in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Psychology

Old Dominion University

April, 1988

Approved by:

Glynn D. Coates (Director)

ABSTRACT

Investigation of Display Issues Relevant to the
Presentation of Aircraft Fault Information

Donald M. Allen
Old Dominion University
Director: Dr. Glynn D. Coates

The present study investigated the effects of different display, hypothesis presentation, information presentation, and parameter presentation styles on pilot performance. It was hypothesized that performance would be maximal using picture-based displays, presenting hypothesis information as a composite, showing only out-of-tolerance parameter information, and when parameter information was displayed as a bargraph. The results of the study indicated that pilot performance was best when employing picture- and text-based displays, when fault hypotheses were displayed as composites. There were no differences in response times when picture- and text-based displays were compared. Subjects' performances were best when hypotheses were displayed as composites compared to when the individual hypotheses were displayed. The display of the out-of-tolerance parameter information resulted in faster overall pilot performances compared to the presentation of all relevant parameter information. There were no differences in performance when parameter information was presented as a bargraph compared to when it was presented as a numerical value. The findings are discussed and recommendations for future research are presented.

Acknowledgments

I would like to express my appreciation to Dr. Glynn D. Coates for his guidance and tolerance during the completion of this work and during my tenure at this university. His foresightedness and willingness to explore new avenues of psychological research greatly contributed to this research and to my educational growth. I would also like to thank the other members of my committee: Kathy Abbott, Dr. Raymond Kirby, and Dr. Janis Sanchez-Hucles. Each of them share the curiosity, technical expertise, and tolerance of ambiguity necessary to teach. Their assistance in the completion of this research and my education is deeply appreciated.

I would like to thank the individuals who voluntarily participated as subjects in this research and to the Vehicle-Operations Research Branch at NASA-Langley. The yearn for knowledge and the improvement of the aircraft industry made this research possible. A special thanks goes out to the pilots of the DET 1 1402 MAS of the United States Air Force who made a special effort to participate in the investigation.

A special expression of appreciation goes to my wife, Stacy, who has tolerated the continual disruptions and ambiguities my life as a student has placed upon her. Her love and understanding contributed to the completion of this dissertation, my education, and my career possible. I can think of no higher praise.

Finally, I would like to thank my fellow graduate students, both past and present, for their support and encouragement. They made life tolerable during the hard times and a joy during the easier times. I hope your dissertations go as well.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

TABLE OF FIGURES

Investigation of Display Issues Relevant to the

Presentation of Aircraft Fault Information

The rapidly growing area of computer technology has begun to produce an increase in the number and type of computer automations found in process control environments (i.e., manufacturing, chemical processing plants, etc). These inroads have become conspicuously evident in the cockpits of high performance military aircraft (Adams, 1987) and to a somewhat lesser degree in the cockpits of commercial air carriers. Evidence of this increase in technology in commercial aircraft is demonstrated by the increase in the number of alerting signals, in the form of warnings, cautions, and advisories that have been incorporated into the designs of the newer aircraft. When the Boeing 707 was introduced in the late 1950's it contained 188 alerting signals, compared to the Boeing 747, introduced in the late 1960's, containing 455 alarms and to the latest class of aircraft which contain over 600 alarms in the form of lights, auditory signals, flags, or bands (Veitengruber, 1977; Hanson, Boucek, Smith, Chikos, Hendrickson, Howison, & Berson, 1982; Boucek, Hanson, Berson, Leffler, & Hendrickson, 1981).

As technology has increased the ability of aircraft designers to monitor various aircraft systems has become more sophisticated, and attempts have been made to convey this information to an already overloaded pilot. The dilemma faced by aircraft designers is to determine from the available information, what is necessary for the pilot to fly the aircraft safely and efficiently without overloading the

pilot? Flight conditions also play a role in this determination. Pilot performance during normal aircraft operation can be improved with more information since he has time available to process the information and to integrate it into his flight operation. However, during abnormal conditions the introduction of information may be more harmful than good, increasing pilot workload and distracting him from the safe operation of the aircraft. The task for the system designer is then one of providing the necessary information in a timely and comprehensible manner while maintaining pilot workload within an acceptable range.

Approaches to Flight Deck Automation

The topic of flight deck automation and the role of the pilot has received a great deal of discussion in the literature. There appears to be agreement among the discussants that some level of automation would improve the performance of flight crews and increase the safety of the passengers and crew. However, there has been little agreement among discussants as to the appropriate types of functions to automate. Three basic approaches have surfaced addressing the types of piloting functions that should be automated and the type of interactions that should occur between the pilot and crew.

On the one hand, Rouse and his colleagues (Chu & Rouse, 1979; Rouse, 1981; Rouse, Rouse, & Hammer, 1982; Rouse, Geddes, & Curry, 1987) working extensively in the automation area, have developed a conceptual design of a pilot/crew support system. Within this framework, the fundamental operating philosophy is one in which all systems are in the

manual control of the flight crew, with automation occurring when the computer anticipates unacceptable performance on the part of the crew or when the crew turns over the control of an operation to the computer. The key theme to this approach is that the pilot is always "in charge" of the aircraft . The pilot has the ability to override the operation of the computer at any time, preserving the crews ability to innovate in situations where the computer would be unfamiliar.

The support system will perform such functions as information management, error monitoring, and adaptive aiding consistent with the "automation as backup" approach. The information monitor manages the presentation of information to the crew based on operator workload, and "need-to-know" priorities. The error monitor assists the pilot to overcome his limitations by identifying and correcting errors, such as errors of omission. Finally, adaptive aiding enhances the operator's performance by using various aiding techniques such as queuing to inform the pilot of the sequence in which necessary tasks need to be completed (Rouse, et. al, 1987). The crew and the support system operate separately, requiring that the pilot monitor the computer and the computer monitor the pilot. While the pilot already possesses the ability to monitor the system, current computing technology is not at the level of sophistication that would permit the computer to monitor accurately the pilot. It is unknown whether people functions in a manner that would permit the mathematical specification necessary to monitor their behavior.

The automation approach of Rouse and colleagues differs from the remaining two approaches in the role that automation plays in the crew-computer interaction. As indicated previously, the role of the computer in Rouse's system is that of a support system, acting only when the crew is not able to function normally. The role of the computer in the remaining two approaches is that of a third crew member (in commercial two- pilot aircraft) or an electronic crew member [in fighter aircraft (Moss & Reising, 1985)] that has its own set of responsibilities and functions independent of the crew. The extra automated crew member is intended to reduce the workload of the pilot by automating some of the responsibilities of the pilot in current aircraft.

The second and third approaches to flight deck automation have their origins in Rasmussens (1983) three theoretical phases of human behavior: skill-based, rule-based, and knowledge-based. Briefly, skill-based behavior is characterized as the recognition of a problem followed by limited data collection, and an immediate conclusion of the appropriate righting action necessary to correct for an abnormal condition. Rule-based behavior demands greater cognition on the part of the system operator, who must apply a set of learned rules to determine the appropriate corrective action. Knowledge-based behavior is necessary under conditions where the operator is faced with unfamiliar situations, where compiled rules do not directly apply, and the operator must formulate hypotheses based on his fundamental knowledge of the system in order to diagnose the system abnormality successfully. The

three types of behavior involve a common sequence of events: abnormality - diagnosis - action. The differences between them involve the level of diagnostic sophistication incorporated in the diagnostic (cognitive) step of the process. Herein lies the difference between the second and third approaches to automation.

Some researchers recommend that circumstances in which response time requirements are near or beyond human capabilities are the most appropriate situations for automation (Weiner, 1984; Hart & Sheridan, 1987). They further suggest that since current technology is not able to support rule-based or knowledge-based behavior, then only those situations that require skill-based behavior should be considered for automation. Hart & Sheridan (1987) argue that the crew should be assigned tasks requiring inductive reasoning, solving unique or novel problems, or learning from experience since the crew might be better at detecting subtle patterns in noisy environment or recognizing/interpreting/processing unpredictable or unexpected events. This position characterizes the second approach to automation -- the crew is responsible for tasks requiring rule-based and knowledge-based behavior, while the system assists in the performance of skill-level tasks where the response times of the pilot are too slow.

The third approach to automation, while based on Rasmussen's behavior taxonomy, takes an opposite view of the situations that are most appropriate for automation. Supporters of this position hold that current technology is at a level that can support knowledge-based

reasoning and that situations requiring this type of response are appropriate for automation (Moss, Reising, & Hudson, 1984; Moss & Reising, 1985). Moss and Reising (1985) argue that pilots are much more likely to trust advice from a system that performs an operation better then they can. Since skill-based and, to some extent, rule-based behaviors are by nature easier for the pilot to perform, it is argued that an automated system would be more acceptable if it addressed knowledge-based activities since these are subject to more uncertainty on the part of the pilot.

There is some scientific evidence that pilot performance is differentially effected by environmental conditions. Such research provides evidence that supports the automation of higher level cognitive tasks. Schudy, Corker, and Baron (1987) have observed that pilots' deep reasoning (or knowledge-based reasoning) in dangerous situations (stress, high pilot workload), given real-time requirements, often led to erroneous short-cuts in problem analysis and ultimately in incorrect pilot response. The environment in commercial airline cockpits during takeoff and landing has been described as being highly stressful and having workload conditions with real-time requirements where aberrant problem solving has the most potential for catastrophe. In a study of pilot operation under stressful conditions, where pilots were asked to attempt to diagnose engine difficulties while flying on autopilot (also faulty), Rockwell, Giffin, & Romer (1983) found that pilots became so engrossed in their attempts at diagnosis they paid no attention to their flight position. This behavior resulted in the majority of the pilots

violating their altitude clearance. Pilots spent from 4 to 10 minutes gathering diagnostic information to the exclusion of other piloting responsibilities, resulting in the unsafe operation of the aircraft. Both studies involved the degradation of pilot performance in situations producing increases in pilot workload that is associated with knowledge-based reasoning.

It is important to be cognizant of the possibility that the introduction of automation in the cockpit will not resolve piloting problems. As indicated in the study discussed above, it was a faulty autopilot paired with the faulty engine that led to the potentially hazardous conditions. In a study investigating the effectiveness of a new autopilot system researchers found that numerous pilot errors could be attributed to the autopilot and more directly to the interface between the pilot and the autopilot (Bergeron, 1981). In a follow-up study pilot errors decreased when adequate effort was extended to improve the pilot-computer interface (Bergeron & Hinton, 1983).

The results of Bergeron's initial study are not surprising. It has become almost commonplace to encounter computer systems that go unused or are improperly used because of a poor interface between the user and the system. This problem becomes more critical as the nature of the interaction between human and computer changes from that of human as operator to that of a human as monitor. This necessarily increases the importance of the human-computer interface as the sole source of system information available to the user (Eberts, Nof, Zimolong, & Salvendy,

1984). This problem manifests itself in the implementation of expert systems where the operator's interaction with the system is most frequent in time-critical situations where system operation is abnormal. The poor presentation of information at this time can paralyze operators, preventing them from recognizing relationships among parameters, recognizing trends, or seeing overall system functionality (Klein & Calderwood, 1986), and severely restricting their ability to return the system to normal operation.

Faultfinder

The present investigation will attempt to identify and to optimize display parameters which influence the effective display of information in a process control environment (piloting commercial aircraft). Specifically, the investigation will be concerned with the display of fault diagnostic information generated by Faultfinder, an expert system under development at the NASA-Langley Research Center (Abbott, Schutte, Palmer, & Ricks, 1987). Faultfinder has been designed to diagnose engine faults for twin turbofan propulsion engines of transport aircraft. Faultfinder will operate in the airplane cockpit of the future that will incorporate five high resolution color cathode ray tubes (CRT), (see Figure 1) with portions of displays dedicated to the presentation of engine, systems, and warning information (see Sexton, 1983, for more details).

Faultfinder will operate as an independent system, following the third approach to automation, where the system will recognize, and

identify engine faults independent of any interaction with the crew. The system gathers the required information from onboard sensors, processes the information, identifies the fault, and presents the fault information to the crew. The interaction between operator and computer differs considerably from that found in previous expert systems such as MYCIN (Shortliffe, 1976) and DENDRAL (Buchanan & Feigenbaum, 1978) that must query the user for information in order to perform any computation. Faultfinder is unique to cockpit automation in that it performs a task (diagnosis) pilots very perform infrequently and for which they receive little training. Rather than limit the fault diagnostic task to a inadequately prepared crew and to increase flight safety, this task can be automated and removed from the pilots list of required tasks during critical phases of flight.

Insert Figure 1 about here

Faultfinder's reasoning is performed using a two-stage process. During Stage 1 the system attempts to identify eight known engine faults (foreign object ingestion, compressor stall, turban blade separation, icing, flameout, anti-surge bleed valve leak, fuel leak, bad fuel control, fuel system failure, and bleed air leak) using traditional expert system approaches, and attempting identification of the most frequently occurring faults first (Abbott, et al, 1987). If the engine parameters do not correspond to one of the eight known faults then Stage 2 processing begins. Stage 2 reasoning is based on knowledge about the

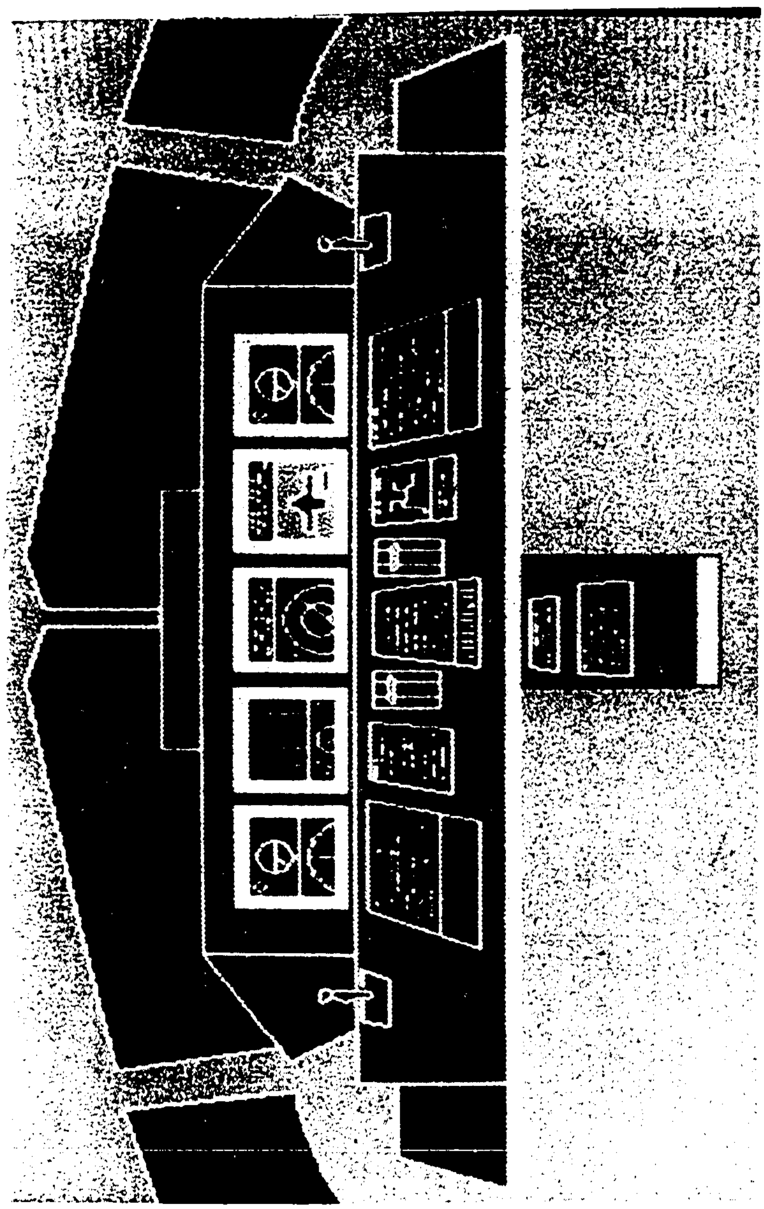


Figure 1. Future Transport Cockpit Display Layout

qualitative models of both the physical and structural components of the engine. Fault identity hypotheses are generated by the system based on current parameter values, and subsequent changes in engine parameters are used to eliminate hypotheses.

Within Faultfinder, faults are classified into levels of abstraction. Stage 1 faults are the most specific, describing a particular malstructure or malfunction, while Stage 2 reasoning, less specific, identifies hypotheses containing fault source components, system status (affected component), and fault propagation path (Abbott, et al, 1987). The information generated by the two stages is quite different, with Stage 1 information being very fault specific and Stage 2 information being less specific. Thus, any display approach intended to represent this information must be able to address both types of information presentation.

Faultfinder currently addresses faults that occur in the engine, however, in its final form it will address all aircraft subsystems. The current investigation included the use of four of the aircraft's subsystems (engine, fuel, hydraulic, and oil).

Experimental Issues

Five display related issues have been identified that are considered relevant to the implementation of the Faultfinder system, as well as to other process control system interfaces. These issues are: 1) the use of color in the interface, 2) whether the interface should be

picture-, symbol-, or text-based, 3) whether the presentation should include fault hypothesis information for the crew, 4) whether parameters are better presented as bargraphs or numbers, and 5) whether the presentation of "surplus" information will improve or hinder performance.

Color. The question of whether or not to include color in CRT displays as it relates to user performance has generated a great deal of research interest; some results have been supportive of its use and others have not. Color is certain to find its way in future cockpit designs since it is used in present day commercial aircraft (L1011, B767). The question of whether to use color or not has changed to one of when is it appropriate to use color and when it is not? Krebs, Wolf, and Sandvig (1978) have identified five situational characteristics that, when present, require color displays to aid operator performance. These were developed as guidelines and are not based on research data. Krebs et. al. believe that color should be incorporated in a display when:

1) the display is unformatted,
2) information presentation is very dense,
3) task requirements include searching for information,
4) there is degraded display legibility, or
5) there exists a logical relationship between the tasks and certain colors.

Many authors and researchers have supported the use of colors

giving justifications for its use such as decreasing workload/scan (Hussey, 1980), improving the integration of multiple electro-optical displays into a single CRT display (Reising & Calhoun, 1982; Kopala, 1979), improving performance when working under time restraints (Benbusat, Dexter, & Todd, 1986), and increasing information differentiation (Jauer & Quinn, 1982). However, there has been some research that has failed to find a difference between operator performance using a color displays compared to black and white displays (Christ, 1977, Kellogg, Kennedy, & Woodruff, 1981; Way, Hornsby, Gilmor, Edwards, & Hobbs, 1983; Tullis, 1981; Reising & Artez, 1983). When asked their preference for colored versus black and white displays, subjects overwhelmingly preferred the color displays (Tullis, 1981; Reising & Artez, 1983, Way et. al., 1983). Given the lack of evidence demonstrating overwhelming performance differences in color versus black and white display, the subjective preference for color, and evidence indicating that high workload and instrument integration tasks benefit from color, it is expected that the airline industry will select color CRT's over black and white for next generation commercial aircraft.

The visual search requirement and the potential for a high degree of information density in the present study also warrants the use of a color system. The task of a crew using <u>Faultfinder</u> in its current form is to determine the subsystem(s) involved in the abnormal condition identified by the system. This involves searching through the display and locating the subsystem and components that are out of tolerance in an environment. Depending on the findings of this study this display

environment could have a high degree of information density (multiple hypotheses presentation with surplus information). Given this, the present research will concentrate on the use of color in cockpit CRT displays without comparing it to black and white displays.

Display Style. The introduction of CRT's into the cockpit has opened a new avenue of information display in that environment. Initial applications of CRT display technology primarily involved the digitization of conventional optical-mechanical displays and did not attempt to utilize the flexibility afforded by the new technology (Way et. al., 1983). More recent researchers have attempted to incorporate some of the flexibility of the CRT into the display of information. This research has centered around the use of text (information displayed in alphanumeric form) and graphics (information present in a symbol or picture form) in the display of aircraft information. Some researchers support the use of text-based displays, advocating the use of multi-paging functions (presentation of information on different displays with controls to page forward or backwards through them) to display information that cannot be displayed in a single alphanumeric display (Graham & Broomhead, 1975; Schneider & Evevoldson, 1984). In an investigation of the use of graphics (symbol/picture) and textual displays in monitoring flexible manufacturing systems where subjects had to identify and correct system faults, Sharit (1985) found that subjects performed better when given information displayed textually than when the information was displayed graphically or when text and graphics were combined. Stern (1984) found similar results when he presented subjects

with procedural instructions in the form of text, voice, or graphics displays. Results indicated that graphics alone was not as effective form of communication as text.

Recently, the results of some investigations supported the use of symbol and graphic based displays to present information. Dickson, Desantis, and McBride (1986) found that when attempting to identify market trends, subject's performances were better when the information was presented in a bar graph as opposed to a table-based presentation. In a study on display interpretation comparing symbolic and text-based displays, Tullis (1981) found that subjects were able to respond faster with the same level of accuracy with the symbol-based displays than with the text-based displays. Speed and accuracy are, of course, both important aspects of the safe piloting of aircraft.

More sophisticated attempts to present information into picture-based displays have been performed in the flight deck environment of combat aircraft (Jauer & Quinn, 1982; Way, et. al., 1984; Summers & Erickson, 1984; Robinson & Eberts, 1987). The research of Jauer and Quinn (1982) and Way et. al. (1984) has focused on an effort to display pictorially all fighter aircraft subsystem statuses (i.e., engine, hydraulic system, electrical systems, fuel systems, and weapons system) with a restricted use of alphanumerics. This effort has concentrated on formatting issues as well as pilot usability and acceptance of picture-based information presentation. Summers and Erickson (1984) took a somewhat different approach to the use of pictorial information.

The primary system display in this display hierarchy was text-based. To traverse through the set of displays a pilot could use a page advance or page back control. The subsystem displays combined graphics (symbols/pictures) and alphanumerics, since the results of previous research had demonstrated that graphic representation of system functional/system relationships requir- fewer cognitive steps to interpret. On the other hand, higher order and abstract concepts have no graphical correspondence and are therefore best represented alphanumerically.

Hypothesis Presentation Style. Faultfinder is unique to other expert systems involving human computer interaction in that it operates in a real-time, continuous process environment and functions independent of the operator. When the system identifies "out-of-tolerance" system parameters it generates a number of hypotheses that correspond to the currently available information. The system then proceeds to eliminate hypotheses whose characteristics fail to correspond to new input information obtained by the system. The information gathering process continues until a single hypothesis remains or no further information becomes available to eliminate further the remaining hypotheses. Faultfinder, given this process of problem solving, could interact with the crew in three different styles. Firstly, it could operate without the crew knowing that anything is wrong with the aircraft (not informing the crew when a fault is first detected) informing the crew only after it has identified the fault. Secondly, Faultfinder could inform the pilot of an abnormal condition when it first recognizes that the

aircraft was not operating normally. Finally, Faultfinder could alert the crew at any point between the discovery and identification of the fault. Working under the assumption that the pilot should be informed of a fault when it has been detected the question becomes in what form should the information be presented to the pilot? Faultfinder generates hypotheses about the system fault based on the information available to it at that time. Therefore, a multiple hypotheses presentation is one form that the information presented to the pilot can take. The set of potential faults that the system has generated can be presented to the pilot in a single display of multiple hypotheses. It is logical to assume that when the diagnostic system generates a set of fault hypotheses about the current state of the operating system those hypotheses would be related. That is, they would deal with the same subsystem, e.g. either the hydraulic system or the fuel system. Therefore, a second method of presenting the information to the pilot is to develop a composite hypothesis based on the similarities of the set of active hypotheses. In this method, the pilot would be presented a display with a single, composite hypothesis, rather than a display with multiple single hypotheses. Examples of these two types of displays are presented in Appendix A.

The two types of displays will necessarily differ in the information contained in them. The individual hypotheses that make up the multiple hypotheses display would contain individual component fault information since they describe a specific potential system fault. Using the aircraft's oil system as a an example, if a fault occurs in

this system and is detected by <u>Faultfinder</u> a set of hypotheses is generated. Each of the hypotheses contained in the multiple hypothesis display would contain information about a subsystem component responsible for the current state of the system. When a composite hypothesis is displayed, it will not contain the individual component information, but will instead contain a representation of the abnormally functioning subsystem. When <u>Faultfinder</u> detects a fault and then generates a set of fault hypotheses based on the current state of the system, the composite hypothesis would display the subsystem suspected of operating abnormally (oil, hydraulic, electrical, etc.) based on the similarities among the current set of individual hypotheses. For example, in the case of the oil subsystem, some indication would be made to the pilot that there is something wrong with the aircraft's oil subsystem giving no reference to potentially responsible subsystem components (oil pump, lines, valves, etc.).

Comparing pilots performance under the two display conditions will determine if pilots are better able to identify current and future system status when given information concerning every potential fault or when only given nonspecific information concerning the affected system. The results of some research suggests that the use of multiple displays to convey related system information results in an abnormal demand on the operators short-term memory (e.g., Goodsmith, 1981).

<u>Parameter Presentation Style.</u> The introduction of CRT displays eliminated the need for conventional dials and gauges that resulted in

the use electro-optical and analog displays. New information display technologies have been produced that can display information in new innovative ways, such as bargraphs and stars, for monitoring process system status (Peterson, Banks, & Gertman, 1981). This display technology was based on the combination of qualitative and quantitative information because of the mechanical operation of the devices. CRT based data transmission faces fewer restrictions and so is more flexible than previous technology. The problem faced by designers attempting to utilize this new technology is what role will previous technology (in this case quantitative information) play in the design of new cockpit displays. The present study will attempt to determine if the inclusion of quantitative information in a qualitative display improves or hampers performance.

There is some disagreement in the cockpit display field of the role of the crew in determining the level of information detail conveyed in system displays. Some researchers suggest that the aircrews should be given the ability to select/control the level of display detail (Schmit, 1984; Murphy & Mitchell, 1986). Others suggest that too much time is already spent on managing information presentation resulting in increase pilot workload and a decrease in pilot effectiveness (Adam, 1987). The present research will attempt to determine the detrimental or facilitative effect that quantitative information plays in the presentation of system status information.

<u>Information Presentation Style.</u> The display of "surplus"

information concerns the inclusion in the display of information that is not directly related to the task at hand of diagnosing the current aircraft fault. This issue relates to the information processing question of what information is necessary for the pilot to operate the aircraft safely under abnormal conditions. Researchers supporting the design guideline of display by exception (where only information necessary for successful task completion is presented to operators), advocate presenting only the information directly related to the fault (out-of-tolerance only; cf, Wanner, 1982; Murphy & Mitchell, 1986; Schudy, Corker, & Baron, 1987). Others suggest that information that is not abnormal but related to the current situation would be helpful for the pilot in determining the overall system status (Mas, Erickson, & Jordan, 1979; Calhoun & Heron, 1981; Hoecker, 1982). It is logical to assume that the inclusion of more information increases the degree of information density in the display and thus displays containing less information should have quicker associated responses.

The current research will attempt to determine if the inclusion of "surplus" information improves or hampers pilot performance when attempting to determine aircraft subsystem status.

Experimental Expectations

The current research addressed four topic areas considered to be important to the presentation of fault information in the commercial airline cockpit. Faultfinder, an expert aircraft fault diagnostic system, was used as the test bed on which to base the investigation.

The first topic area concerned the style of presentation. Recent research suggested three types of display styles that would be appropriate: picture-, symbol-, and text-based. While previous research has shown that pilots can operate systems with text or graphical based displays, it is not clear under what conditions pilots will perform maximally. Therefore, the present study incorporated an assessment of the picture-, symbol-, and text-display modes. It was expected that pilots would perform best when using the picture-based displays and poorest on the symbol-based displays. It was expected that pictorial displays would be associated with the best pilot performance because they would have more intrinsic meaning associated with them than the symbolic displays, thus requiring less interpretation. Subjects were expected to perform better on the text-based displays than the symbol-based displays since any meaning that becomes associated with a symbol would have to be learned and the training employed in the study would not be sufficient to overcome the subjects' existing reading abilities.

The second topic of research was unique to Faultfinder and relates to the presentation of individual fault hypotheses or the presentation of a composite hypothesis. Pilot's performance and pilot preference are expected to support the use of single composite hypothesis presentation because of the time and effort required to integrate the separate fault hypothesis into a meaningful chunk of information.

The third topic area is concerned with the style of parameter presentation and will compare subject performance using bargraphs and

simple numerical values. It was expected that pilots would prefer displays that included quantitative parameter information; however, their performance was not expected to be any better, and perhaps, even worse, when the information was presented in a bargraph.

The fourth topic area addresses the issue of determining what parameter information is necessary for pilot performance and attempted to identify how information should be selected for inclusion in the displays. An attempt was made to determine whether pilot performance would be different under two styles of information presentation. One information presentation style advocates the presentation of all information relevant to the fault in the displays, while the second suggests that only those parameters that are operating abnormally are appropriate for inclusion in the displays. It was expected that performance would be quicker in situations where only the out-of-tolerance information is presented. It was further expected, however, that pilots would prefer the inclusion of all relevant information, choosing to determine for themselves what is necessary for the safe operation of the aircraft.

Since the different issues addressed in this study have not been investigated simultaneously in previous research, one of the purposes of this investigation was to examine the interactions among them.

Method

The present investigation compared the performance of subjects as a function of different techniques for presenting system status information to pilots based on the information that might be generated by Faultfinder. Subjects were asked to determine current system status using displays generated according to the issues discussed above.

Design

The experimental design was a 8 Case by 3 (Display Style) by 2 (Hypothesis Presentation Style by 2 (Information Presentation Style) by 2 (Parameter Presentation Style) by 4 (Pilot Type) model. All variables were within-subjects measures. Subjects' response times and accuracy were used to compare the effectiveness of the different display variables.

Subjects

Subjects were 18 male pilots each with a minimum of 500 hours of flight experience in twin turbo-fan engine aircraft. The range of pilot flight experience was 500 to 22,000 hours of overall flight experience. The subjects were groups of six NASA test pilots, six pilots from two major commercial airlines, and six U.S. Air Force pilots. Six male non-pilots, who were familiar with flight deck displays and aircraft operation, were also tested. The inclusion of non-pilots in the subject pool was intended to assist in determining pilot performance that is a

product of previous training and not the result of experimental displays, and since it is unclear how receptive some of the subjects would be to information displays that differ from those to which they were accustomed. Each subject participated in all experimental conditions.

Apparatus

The displays were presented on a Casper EGA (640 X 350) color monitor, model TE5154, driven by a Gem G/AT computer running at 10MHZ. Subjects' response times were recorded by the computer. A Panasonic portable tape recorder was used to record subjects' verbalizations.

Stimulus Material

The displays used in the investigation were created using Dr. Halo II, a rapid prototyping tool using a 16-color palette and fat-bits editing capability. The displays were generated based on accident reports obtained from the National Transportation and Safety Board describing commercial airline accidents involving multiple turbo-fan aircraft that had occurred between 1970 and 1986 as well as incident reports obtained from the NASA Aviation Safety Reporting System. Information concerning the eight test cases employed in the study are presented in Table 1. The cases are ordered in terms of the amount of information presented in the display in both the composite and multiple hypotheses conditions. The cases fall into three logical groups in terms of complexity. The first group is made up of Cases A, B, C, D, and E which involve faults occurring in one of the four subsystems. The

second group, containing Cases F and G involve 2 of the four subsystems, while the last group, Case H, involves three of subsystems. The cases were selected to represent the full range of faults that a pilot might logically expect to encounter during the operation of the aircraft.

Insert Table 1 about here

The cases selected for inclusion in the investigation represented accidents involving single and multiple aircraft subsystem failure. The selected cases included situations requiring Stage 1 and Stage 2 reasoning as dictated by Faultfinder. Five of the test cases involved one hypothesis in the multiple hypotheses condition, two involved two hypotheses, and the remaining case involved four hypotheses. Each test case was represented by 24 displays generated to represent the different combinations of research issues (Composite vs Multiple Hypotheses X Picture- vs Symbolic- vs Text-based Displays X Bargraph vs Numerical X All Relevant vs Out-of-tolerance Only). A set of representative displays for the eight test cases is presented in Appendix A.

The group of displays were presented in three separate blocks based on the content of the display: text-based, symbol-based, picture-based. Each block contained 64 displays (Multiple vs Composite Hypotheses X Bargraph vs Numerical Value X All Relevant vs Out-of-tolerance Only X 8 Test Cases). The major display blocks were separated into two groups: multiple hypotheses and composite hypothesis displays. The presentation

Table 1

Experimental Test Cases

Case A

This was a hypothetical case involving a right hydraulic system failure. A low hydraulic pressure indication was the only out-of-tolerance parameter. The multiple hypotheses display contained a single individual hypothesis indicating a failed hydraulic pump.

Case B

This cased was based on NTSB report number NTSB/AAR-84/04 involving a Lockheed L1011 operating as Eastern Flight 855. The case involved a left oil system failure, with oil quantity and oil pressure both low. The multiple hypothesis display contained a single individual hypothesis depicting an oil leak.

Case C

This was a hypothetical case that involved the right fuel system. Fuel flow and fuel pressure were both low. The multiple hypothesis display contained a single hypothesis depicting a fuel leak.

Case D

This was a hypothetical case involving the left engine. A low N1 sensor reading was the only out-of-tolerance information. The multiple hypothesis display contained two individual hypotheses, one depicting a failed fan, and the other depicting a failed N1 sensor.

Case E

This was an accident case based on National Transportation and Safety Bureau (NTSB) Report Number NTSB-aar-76-19 involving Douglas DC-10-30 operating as Overseas National Airways Fight 032. The case involved a single left engine failure with N1, N2, EPR, and EGT all low. The multiple hypotheses displays contained four individual hypotheses indicating a failed fan, compressor, combustor, and turbine separately.

Table 1 (continued)

Experimental Test Cases

Case F

This case was also based on NTSB report number NTSB/AAR-84/04, but represented the failure in a later phase of its development. The left engine as well as the left oils system were involved, with the oil quantity, oil pressure, N1, N2, EGT, and EPR all low. The multiple hypothesis display contained a single hypothesis depicting an oil leak and a failed fan, compressor, combustor, and turbine.

Case G

This was a hypothetical case involving a right engine and right fuel system failure. N1, fuel flow, fuel pressure and fuel quantity were all low. The multiple hypotheses displays contained two hypotheses, one involving a failed fan and fuel leak and the other a failed N1 sensor and fuel leak.

Case H

This case was based on NTSB accident report number NTSB-AAR-82-5 involving a Lockheed L-1011-384 operating as Eastern Airlines Flight 935. The case involved failure in the right engine, right oil system, and right hydraulic system. N1, oil pressure, and hydraulic pressure were all low. The multiple hypotheses display contained a single hypothesis depicting a failed fan, an oil and hydraulic leak.
of the separate display blocks and the different display groups were partially counterbalanced. Each display group was composed of 32 displays which were randomized for each subject.

partially counterbalanced. Each display group was composed of 32 displays that were randomized for each subject.

Color. Color was used in the presentation of the text-, symbol-, and picture-based information to indicate the operational status of the subsystem or subsystem's components. The basic approach to the use of color in display generation was to use grays to represent background information, black and white for outlining and for alphanumerics, red to denote abnormal conditions, and green to denote normal operation. Color was also used to identify leaks in the oil (black), fuel (orange), and hydraulic (magenta) subsystems in picture-based displays and as supplemental data for leaks in the symbol-based displays.

Numerical Parameter Information. Numerical parameter information was displayed numerically as a digital readout in black, with background color indicating operating status (normal-green, abnormal-red). An arrowhead pointing up or down was used to indicate whether an abnormal parameter was abnormally-high or abnormally-low, respectively.

Bargraph Parameter Information. Bargraph parameter information was displayed using a bar gauge designed for this study (see Figure 2). The green area of the bar represented normal operation, while the red indicated abnormally-high or abnormally-low operation. The yellow areas of the bars were included to promote realism and were not used in the study. A floating horizontal white bar indicated the status of the component.

Insert Figure 2 about here

Instructions and Training

Each display block was preceded by a training session where subjects received instruction on the particular experimental display conditions they were to see and then observed twelve example displays. Most subjects required a single training session, however, two subjects felt they needed to work through one of the sessions a second time. Subjects were asked the comprehension questions during their first training session only.

Text-based displays. During this phase of the study, subjects received instruction and training on the text-based displays which described them as being primarily alphanumerically based. The training included a written description of the displays as well as an example of the display.

Symbol-based Displays. Subjects instruction and training included a written description of the set of displays as well as pictures of the set of symbols that were to be depicted in the displays. The symbols used were geometric shapes that represented the four subsystems included in the study. The symbols used in this set of displays can be found in Appendix B. One symbol was used to represent each of the subsystems.

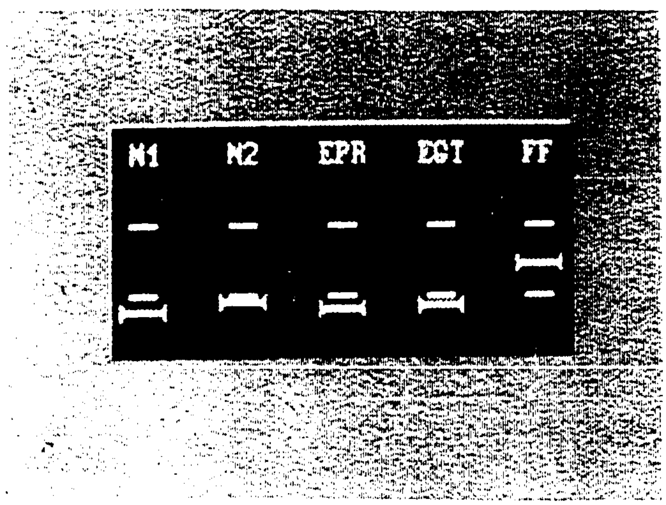


Figure 2. Example of Bargraph Parameter Presentation Display

In the multiple hypotheses displays involving the fuel, hydraulic and oil systems, the subsystem symbol contained a cross to represent a pump failure or was filled to represent a leak.

Picture-based Displays. With this set of displays subjects received instruction and training that included a written description as well as pictures of the display elements contained in the set of displays. The set of display elements used in display construction are located in Appendix C.

Procedure

Subjects were given a brief description of the purpose of the research project that stressed the experimental nature of the system and the importance of considering the type of information they would prefer to have available to them in the flight deck, unrestricted by what is currently available. Subjects were instructed to consider each display in the context of normal flight operation and weather, while flying at a normal cruising altitude. The imaginary plane they would be flying was describe as a next generation aircraft that would have twin turbo-fan engines and would operate in a primarily CRT-based environment. The display they would see would appear on a secondary display panel and would serve as supplemental information to other status information they would expect to have available in the cockpit. A brief description and purpose of Faultfinder was also given at this time. Both written and verbal instructions were given about the experimental process and the tasks they were asked to perform. A copy of the written instructions

given to the subjects is presented in Appendix D.

Task

For each of the presented displays (i.e., 192) the subject's task was to identify, from the display, the aircraft subsystem(s) that had been represented as operating abnormally. The task involved identifying the number and type of subsystems presented in the display and reporting them to the experimenter. They were instructed to work as quickly as possible, but not to sacrifice speed for accuracy.

During the presentation of the displays, subjects were occasionally asked to indicate the operating status of a parameter that was contained in the display the subject had just viewed. The subjects' task was to indicate whether the parameter was abnormally-high, abnormally-low, or normal. This question was asked an average of 35 times per subject.

After completing the experiment, subjects were asked to complete a followup questionnaire that asked them to indicate their display preference and to make suggestions that they felt might improve the set of displays employed in the study. A copy of the post-experiment questionnaire is represented in Appendix E.

Results

Response Time

Response time obtained as a function of test case (A to H), presentation style (picture-, symbol-, and text-based), hypothesis presentation (composite and multiple), parameter information (bargraphs and numerical values), information density (out-of-tolerance only and all relevant), and pilot type (NASA-test, Air Force, Commercial, and Non-Pilot). Subjects were nested in pilot type. Response time data were analyzed using a six-factor analysis of variance which is summarized in Table 2.

Insert Table 2 about here

Main Effects

As can be seen from Table 2, three significant main effects were obtained from the response time data.

Cases. A main effect for cases was found for response time. The mean reaction times for the eight test cases are presented in Figure 3. The differences in the mean response times for individual cases are consistent with what would be expected from the information contained in Table 1. A Tukey (hsd) post hoc test demonstrated 4 small group clusters with differing response times similar to the information complexity groupings described previously. The first grouping was

Table 2

Analysis of Variance Summary Table for Response Time Data

Source	df	Mean Square	F
Case (C)	7	1585620946.00	78.12*
Display Style (DS)	2	28354456.56	<1
Hypothesis Style (HS)	1	812120564.26	49.80*
Information Style (IS)	1	2048600043.95	85.90*
Parameter Style (PS)	1	5831823.98	<1
Pilot Type (P)	3	49152027.28	<1
Subject(Pilot Type) [S(P)]	20	56283353.46	
C X DS	14	8943773.35	1.69
C X HS	7	43014683.31	11.81*
C X IS	7	1152894734.00	22.42*
C X PS	7	3200356.07	<1
C X P	21	10284835.41	<1
C X S(P)	140	2029804.98	
DS X HS	2	66413199.03	3.93*
DS X IS	2	1918971.55	<1
DS X PS	2	1739207.71	<1
DS X P	6	28354456.56	<1
DS X S(P)	40	5084468.38	
HS X IS	1	472675.54	<1
HS X PS	1	4901710.96	<1
HS X P	3	2158089.26	<1
HS X S(P)	20	16306127.07	
IS X PS	1	1064887.70	<1
IS X P	3	11999866.98	<1
IS X S(P)	20	23847306.32	
PS X P	3	5408446.09	<1
PS X S(P)	20	5694534.73	
C X DS X HS	14	6230481.17	1.83*
C X DS X IS	14	3567419.12	1.07
C X DS X PS	14	5102521.39	1.57
C X DS X P	42	2988237.76	<1
C X DS X S(P)	280	5654600.04	
C X HS X IS	7	2407507.49	<1
C X HS X PS	7	940438.74	<1
C X HS X P	21	5086130.20	1.49
C X HS X S(P)	140	3424124.26	
C X IS X PS	7	7205826.39	1.89
C X IS X P	21	4236531.53	<1
C X IS X S(P)	140	5740909.15	
C X PS X P	21	4114337.82	1.20
C X PS X S(P)	140	3431280.91	

Table 2 (continued)

Analysis of Variance Summary Table for Response Time Data

Source	df	Mean Square	F
DS X HS X IS	2	385606.14	<1
DS X HS X PS	2	1521462.26	<1
DS X HS X P	6	24117280.10	1.43
DS X HS X S(P)	40	16884799.28	
DS X IS X PS	2	1952639.13	<1
DS X IS X P	6	4303338.68	<1
DS X IS X S(P)	40	6433796.14	
DS X PS X P	6	4025134.67	<1
DS X PS X S(P)	40	5868375.61	
HS X IS X PS	1	2846597.42	<1
HS X IS X P	3	4032105.06	<1
HS X IS X S(P)	20	6037318.41	
HS X PS X P	3	16920683.76	1.57
HS X PS X S(P)	20	10758774.10	
IS X PS X P	3	7516092.79	1.66
IS X PS X S(P)	20	4516985.06	
C X DS X HS X IS	14	3919488.38	1.30
C X DS X HS X PS	14	2669174.12	1.16
C X DS X HS X P	42	3740222.92	1.10
C X DS X HS X S(P)	280	3404793.75	
C X DS X IS X PS	14	3450531.39	1.16
C X DS X IS X P	42	3517809.03	1.06
C X DS X IS X S(P)	280	3316635.41	
C X DS X PS X P	42	3873298.53	1.22
C X DS X PS X S(P)	280	3169532.08	
C X HS X IS X PS	7	3239924.20	1.02
C X HS X IS X P	21	2033113.13	<1
C X HS X IS X S(P)	140	3275875.84	
C X HS X PS X P	3	16920683.76	1.57
C X HS X PS X S(P)	140	3104746.04	
C X IS X PS X P	21	6195737.89	1.79*
C X IS X PS X S(P)	140	3454931.00	
DS X HS X IS X PS	2	1089023.98	<1
DS X HS X IS X P	6	1543812.74	<1
DS X HS X IS X S(P)	40	2281948.46	
DS X HS X PS X P	6	1077540.97	<1
DS X HS X PS X S(P)	40	6012963.35	
DS X IS X PS X P	6	1968848.21	<1
DS X IS X PS X S(P)	40	4877613.31	
HS X IS X PS X P	3	10032435.38	2.37
HS X IS X PS X S(P)	20	4224245.47	

Table 2 (continued)

Analysis of Variance Summary Table for Response Time Data

Source	df	Mean Square	F
C X DS X HS X IS X PS	14	4878436.33	1.69
C X DS X HS X IS X P	42	3091094.42	1.02
C X DS X HS X IS X S(P)	280	3016266.80	
C X DS X HS X PS X P	42	2307099.22	1.00
C X DS X HS X PS X S(P)	280	2305269.31	
C X DS X IS X PS X P	42	2606115.64	<1
C X DS X IS X PS X S(P)	280	3016315.37	
C X HS X IS X PS X P	21	2598484.72	<1
C X HS X IS X PS X S(P)	140	3265500.78	
DS X HS X IS X PS X P	6	1673206.91	<1
DS X HS X IS X PS X S(P)	40	516118.34	
C X DS X HS X IS X PS X P	42	2785852.82	<1
C X DS X HS X IS X PS X S(P)	280	2907819.52	

* $p < .05$

composed of Cases A (3.520 sec), B (3.459 sec), and D (3.217 sec) with response times 3.520, 3.459, and 3.217 seconds, respectively. The second grouping was composed of Cases C and E with response times of 4.399 and 4.506 seconds, respectively. The third grouping was made up of Cases F and G with response times of 6.269 and 6.684 seconds, respectively. The fourth grouping contained a single Case H with a mean response time of 7.548 seconds. Response times to the first grouping were the quickest followed by the second, third, and fourth groupings, in that order. These results are consistent with what would be expected by the amount of complexity found in the different cases, with Case H having portrayed the greatest amount of information and Cases A and B the least amount of information.

Insert Figure 3 about here

Hypothesis Presentation Style. The significant effect for hypothesis presentation style indicates that subjects were able to respond faster to composite hypothesis displays than multiple hypothesis displays, with mean response times of 4.530 and 5.370 seconds, respectively (See Figure 4).

Insert Figure 4 about here

Information Presentation Style. The data presented in Figure 5

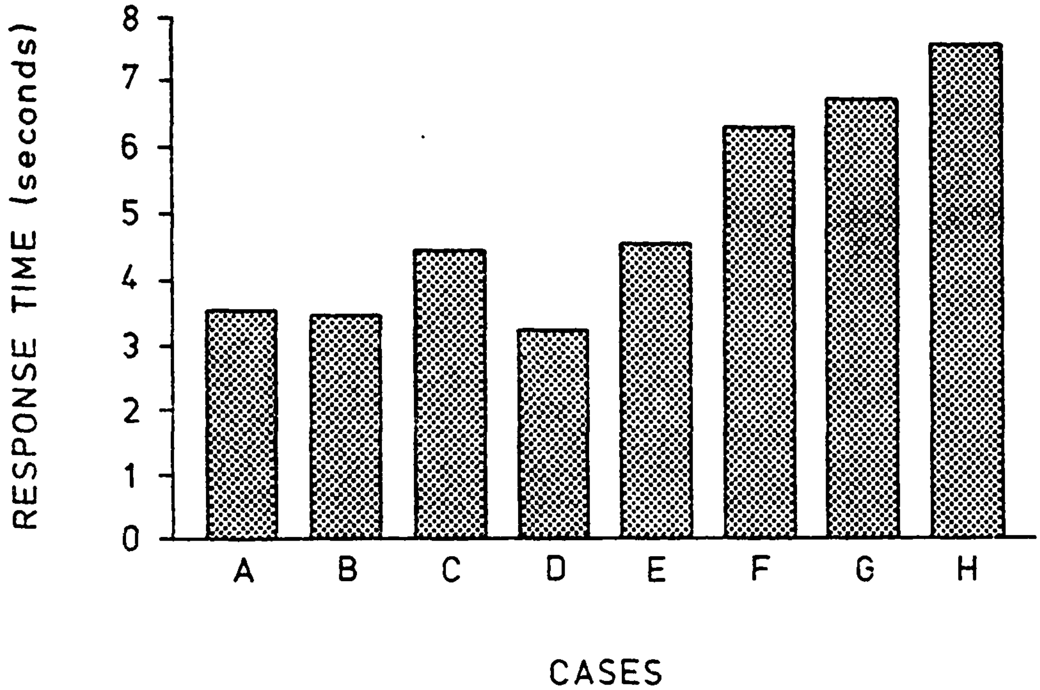


Figure 3. Response Times for Cases.

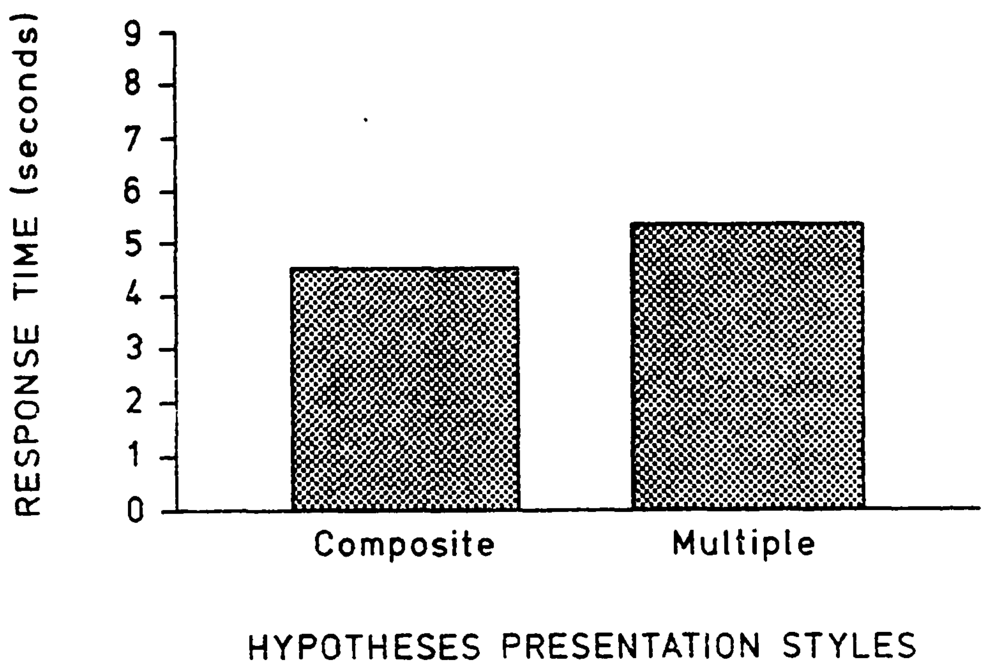


Figure 4. Response Time for Hypothesis Presentation Style

indicated that response times were faster overall in the "out-of-tolerance only" information presentation condition (4.283 sec) than they were in the "all relevant" condition (5.617 sec). There were no differences in accuracy between the two conditions.

Insert Figure 5 about here

Case by Hypothesis Presentation Effects

A significant hypothesis presentation method by case interaction was found. As can be seen in Figure 6, response time was quickest, overall, in the composite hypothesis condition, a Tukey (hsd) analysis confirmed this observation out in all cases except Cases B and C, where no differences were observed. Groupings similar to those found in the main effect for case (discussed above) and the case by display style by hypothesis presentation style case interaction (discussed below) were found when response time for the composite and multiple hypothesis conditions were observed separately. For the composite hypothesis condition the Case F (5.960 sec) and G (5.816 sec) grouping was found, as was the Case C (4.187 sec) and E (3.711 sec) grouping, and an Case A (3.247 sec), B (3.354 sec) and D (7.101 sec) grouping. In the multiple hypotheses condition, the Case A (3.792 sec), B (3.562 sec), and D (3.570) group was the only one observed.

Case by Information Presentation Style Effects

A significant case by information presentation style interaction

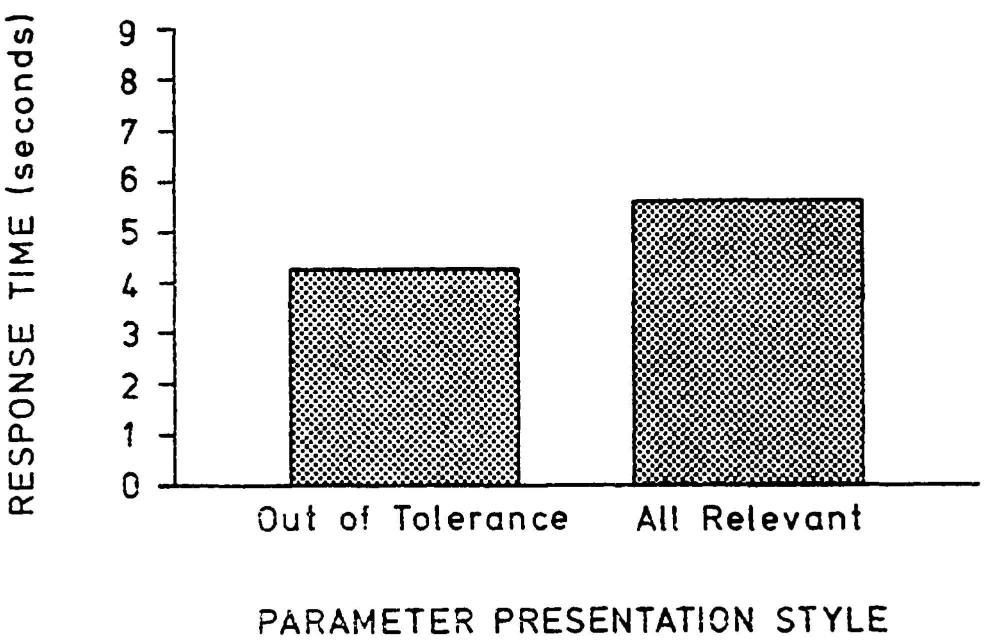


Figure 5. Response Time for Parameter Presentation Style

was obtained from the response time data. The group means are presented in Figure 7. The Tukey (hsd) analysis revealed that subjects' response time was consistently faster in the "only out-of-tolerance" condition of information density when compared to the "all relevant" condition.

Insert Figure 6 about here

Analysis of cases within the "out-of-tolerance" condition revealed a similarity in responses for Cases A (3.106 sec), B (3.105 sec), and D (2.747 sec), for Cases C (3.916 sec) and E (4.071 sec), and for Cases F (5.895 sec) and G (5.125 sec). As was found in the main effect for cases, the more information contained in the displays the longer subjects' took in making their response, and displays with approximate amounts of information had similar response times.

In the "all relevant" analysis the first two complexity groupings were also found, however, Case H replaced Case G in the third grouping.

Insert Figure 7 about here

Case by Display Style by Hypotheses Presentation Style Effects

A significant case by display style by hypothesis presentation style interaction was obtained for response time. As can be seen in

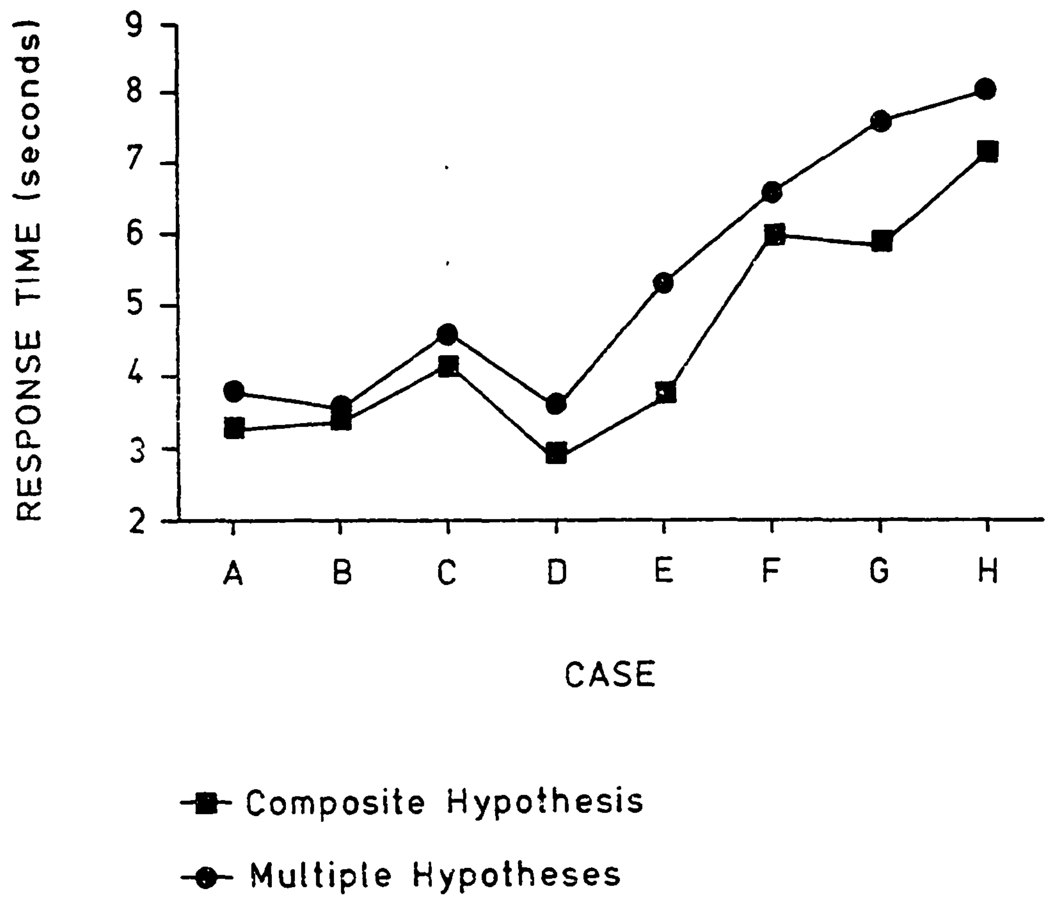


Figure 6. Response Time for Case by Hypothesis Presentation Interaction

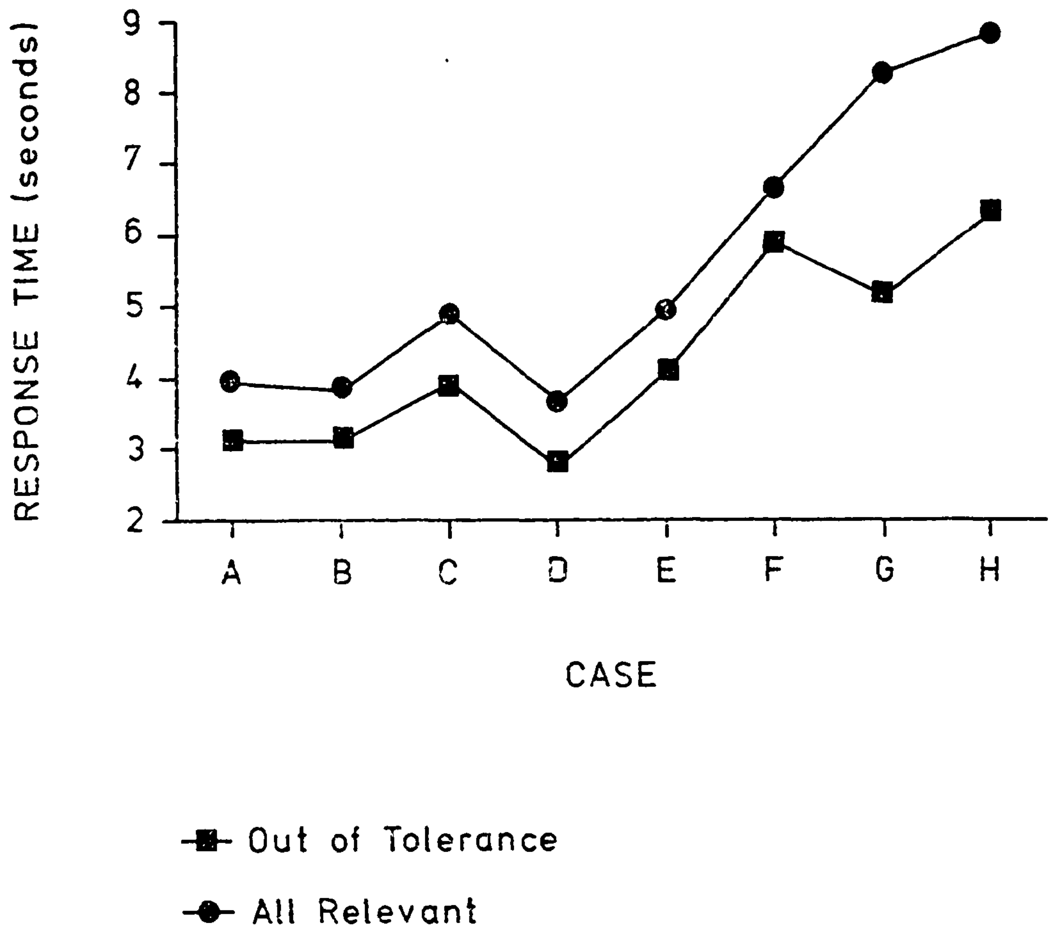


Figure 7. Response Time for Case by Information Presentation Interaction

Figure 8, the effect of hypothesis presentation method on response time was inconsistent across display style and case. These findings suggest that display style and hypothesis presentation style differentially effect subject response time as a function of amount of information presentation. When the amount of information contained in a display is small there are no differences in subjects' response times. However, as this amount increases response times increase for the multiple hypothesis displays as well as the symbol-based displays, with the picture/composite and text/composite displays having the quickest response times.

For Cases A, B, C, and D Tukey's (hsd) analysis revealed no differences between response time across the different Presentation Styles and Hypothesis Presentation Styles. For Case H the differences were found between the picture-composite condition (6.317 sec) and picture-multiple condition (8.167 sec), picture-composite and both symbol conditions [composite (7.692 sec), multiple (7.968 sec)], and the picture-composite and the text-multiple condition (7.846 sec). The findings for Case F were the same as for Case H, except the differences between picture-composite (5.151 sec) and picture-multiple (6.279 sec) and the picture-composite and symbol-composite condition (6.701 sec) were not found.

Insert Figure 8 about here

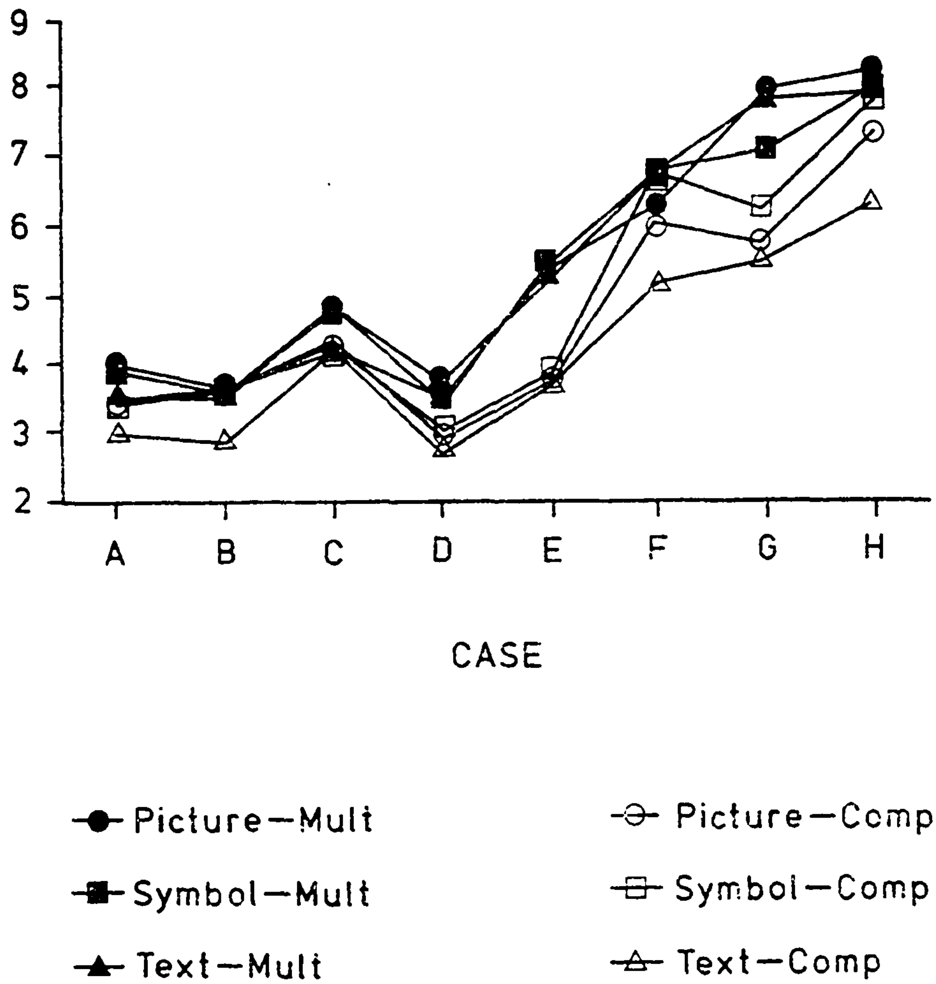


Figure 8. Response Times for Case by Display Style by Hypothesis Presentation

For Case E differences between the picture-composite conditions (3.632 and 5.476 sec, respectively) and the picture- (5.330 and 7.843 sec), symbol- (5.375 and 7.030 sec), and text-multiple (5.194 and 7.780 sec) conditions were found as were the differences between the picture-multiple condition and the symbol- (3.816 and 6.246 sec) and text-composite (3.685 and 5.726 sec) conditions, the symbol-multiple condition and the text-composite condition, and the text-composite condition and the text-multiple condition. Similar results were found for case G except that the difference between the symbol-multiple condition and the text-composite condition found for Case E, was not found for Case G.

Analysis of differences across cases indicated three groupings of cases where subjects' response times were consistent across the six conditions. The first grouping was composed of Cases C and E, the second grouping of Cases F, G, and H, and the third grouping Cases A, B, C, and D.

Case by Information Presentation Style by Parameter Presentation Style by Pilot Type Effect

A significant case by information presentation style by parameter presentation style by pilot interaction was obtained from subject response time. The mean response times are presented in Figures 9 and 10. Trends consistent with those found in earlier analysis were observed. Comparing subjects' response times across the information presentation conditions suggests that when the information presented in

the displays pertained to a single subsystem (Cases A, B, C, D, and E) there were minor differences in subjects' response times. However, when displays contained information about faults in two subsystems (Cases F, G, and H) response times were consistently faster when only the out-of-tolerance parameter information was presented compared to when all relevant parameter information was presented.

Insert Figures 9 and 10 about here

The performance of the different pilot groups did not exhibit any readily observable patterns or trends.

Display Style by Hypothesis Presentation Style

A significant display style by hypothesis presentation style interaction was found for response time. As can be seen in Figure 11 response time for the composite displays (picture 4.153 sec, symbol 4.824 sec and text 4.612 sec) were consistently lower than those for the multiple hypothesis displays (5.440, 5.288, and 5.381 sec, respectively). A Tukey (hsd) analysis confirmed this for the picture and text displays, however, no difference was found between the symbol-composite and the symbol-multiple conditions. Analysis of differences across the hypothesis presentation condition revealed that their was no difference in the response times for the text- and picture-composite displays, but both were significantly faster than response times in the symbol-composite condition. No differences were found in the response

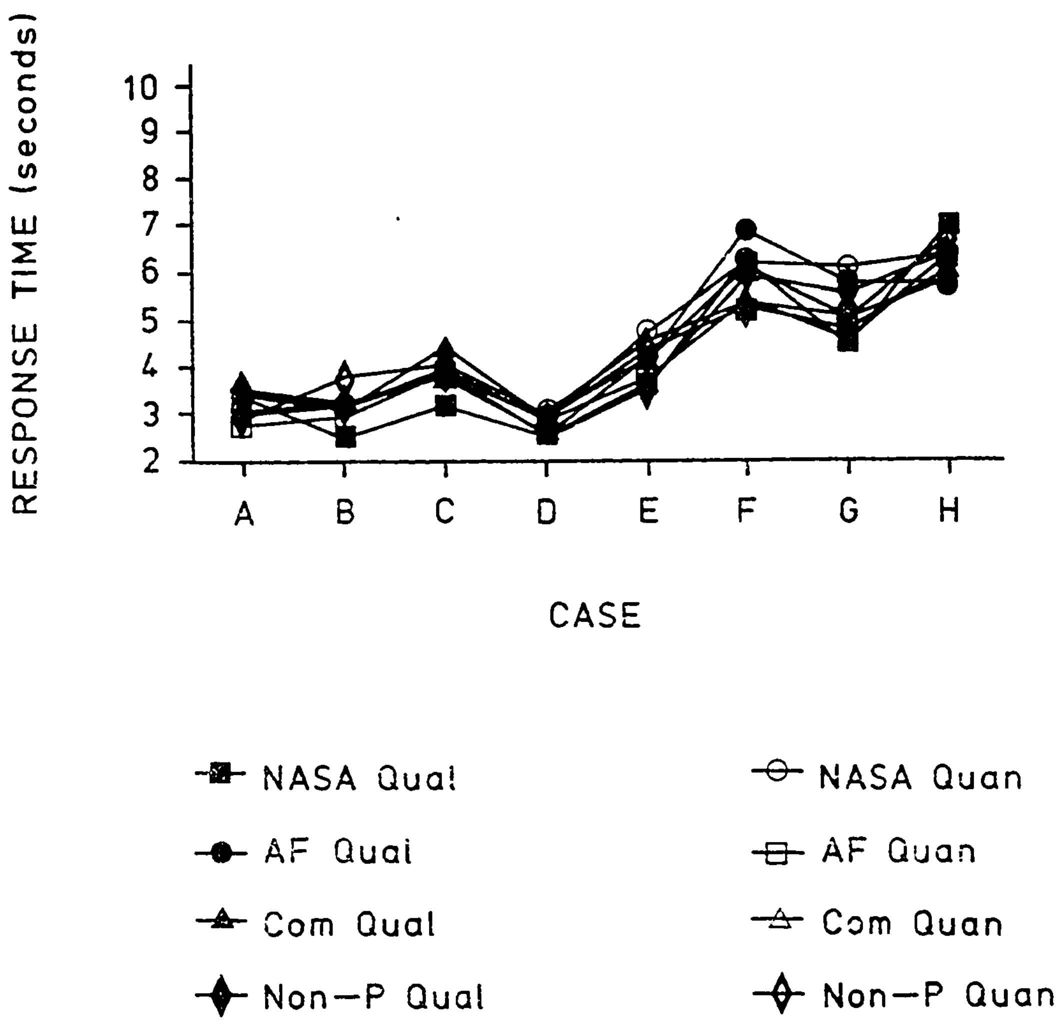


Figure 9. Response Times for Pilots, Cases, and Hypothesis Presentation Interaction —Out of Tolerance Only

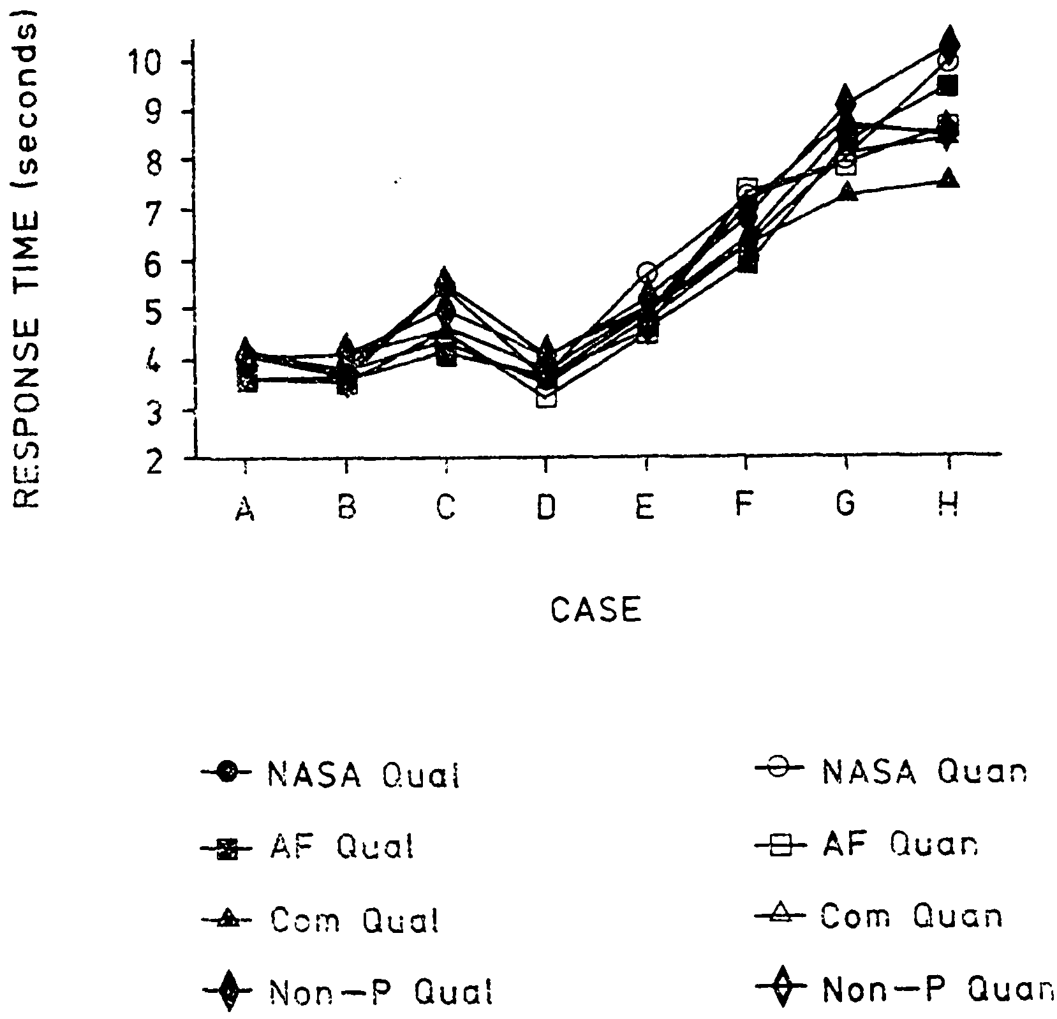


Figure 10. Response Times for Case, Hypothesis Presentation and Pilot Interaction—All Relevant

times for the display styles in the multiple hypothesis condition.

Insert Figure 11 about here

Response Accuracy

Subject response accuracy was separated into number of errors made when identifying subsystems represented in the displays (subsystem error) and percent error made when identifying the operating status of a parameter indicated by the experimenter (parameter error). Subsystem error was calculated by dividing the total number of errors subjects made in each display block and dividing that number by 32 (i.e., the of displays per block).

Parameter error was calculated based on the number of occasions the subject was asked to make this identification per display block. The number of times subjects were required to identify the operating status of a specified parameter varied. The number of times subject's preformed the task ranged from 33-38, with a mean of 35.67. Overall subjects' accuracy was 96.1% for subsystem identification and 95.7% for parameter identification. Display and parameter error were analyzed using separate five-factor analyses of variance, the results of which are presented in Tables 3 and 4, respectively.

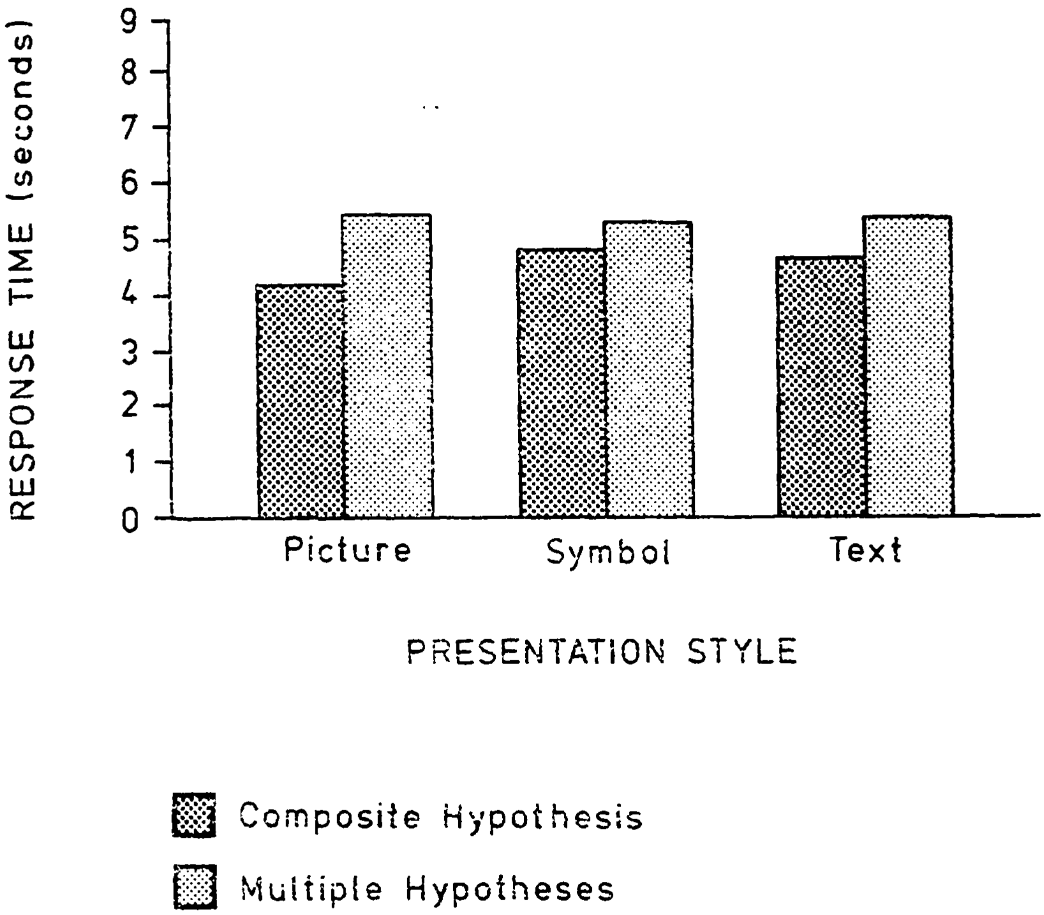


Figure 11. Response Times for Display Style by Hypothesis Presentation Interaction

Insert Table 3 and 4 about here

Main Effects

As can be seen from Tables 3 and 4 one significant main effect was found for subsystem error.

Hypothesis Presentation Style. A significant effect of hypothesis presentation style on subsystem error was found; the means, presented in Figure 12, indicate that subjects made fewer errors in the composite display conditions (.10 per 96 displays) than in the multiple hypotheses display condition (.17 per 96 displays).

Insert Figure 12 about here

Display Style by Hypothesis Presentation Style Effects

A significant interaction was obtained for the display style and hypothesis presentation interaction for subsystem error. As can be seen in Figure 13, and as a Tukey (hsd) confirmed, the number of system errors made in the picture-multiple (0.250) condition were significantly greater than those made in the picture-composite (0.042) condition. No differences were found for the symbol- and text-based displays.

Table 3

Analysis of Variance Summary Table for Subsystem Error Data

Source	df	Mean Square	F
Display Style (DS)	3	0.0365	<1
Hypothesis Style (HS)	1	0.7656	10.40*
Information Style (IS)	1	0.0017	<1
Parameter Style (PS)	1	0.0434	<1
Pilot (P)	3	0.2378	<1
Subject(Pilot Type) [S(P)]	20	0.3427	
DS X HS	2	0.6615	3.54*
DS X IS	2	0.0853	<1
DS X PS	2	0.0851	<1
DS X P	6	0.0920	<1
DS X S(P)	40	0.2073	
HS X IS	1	0.0851	<1
HS X PS	1	0.5017	4.14
HS X P	3	0.1036	1.50
HS X S(P)	20	0.0695	
IS X PS	1	0.0156	<1
IS X P	3	0.0990	<1
IS X S(P)	20	0.1497	
PS X P	3	0.1591	1.64
PS X S(P)	20	0.0969	
DS X HS X IS	2	0.2622	1.79
DS X HS X PS	2	0.2726	1.47
DS X HS X P	6	0.1522	<1
DS X HS X S(P)	40	0.1920	
DS X IS X PS	2	0.1927	1.76
DS X IS X P	6	0.1788	1.32
DS X IS X S(P)	40	0.1351	
DS X PS X P	6	0.1175	<1
DS X PS X S(P)	40	0.1385	
HS X IS X PS	1	0.2101	4.47*
HS X IS X P	3	0.2286	1.94
HS X IS X S(P)	20	0.1177	
HS X PS X P	3	0.0341	<1
HS X PS X S(P)	20	0.1344	
IS X PS X P	3	0.1036	<1
IS X PS X S(P)	20	0.2566	
DS X HS X IS X PS	2	0.1580	1.07
DS X HS X IS X P	6	0.1001	<1
DS X HS X IS X S(P)	40	0.1531	
DS X HS X PS X P	6	0.3050	1.82
DS X HS X PS X S(P)	40	0.1677	

Table 3 (continued)

Analysis of Variance Summary Table for Subsystem Error Data

Source	df	Mean Square	F
DS X IS X PS X P	6	0.1835	1.87
DS X IS X PS X S(P)	40	0.0983	
HS X IS X PS X P	3	0.0480	1.02
HS X IS X PS X S(P)	20	0.0469	
DS X HS X IS X PS X P	6	0.0376	<1
DS X HS X IS X PS X S(P)	40	0.1635	

* $p < .05$

Table 4

Analysis of Variance Summary Table for Sensor Error Data

Source	df	Mean Square	F
Display Style (DS)	3	0.0365	<1
Hypothesis Style (HS)	1	0.0069	<1
Information Style (IS)	1	0.1736	1.53
Parameter Style (PS)	1	0.0069	<1
Pilot Type (P)	3	0.0694	<1
Subject(Pilot Type) [S(P)]	20	0.1021	
DS X HS	2	0.0122	<1
DS X IS	2	0.0018	<1
DS X PS	2	0.1372	1.97
DS X P	6	0.1059	1.52
DS X S(P)	40	0.0698	
HS X IS	1	0.0000	<1
HS X PS	1	0.0278	<1
HS X P	3	0.0671	1.70
HS X S(P)	20	0.0396	
IS X PS	1	0.0278	<1
IS X P	3	0.0208	<1
IS X S(P)	20	0.1132	
PS X P	3	0.0116	<1
PS X S(P)	20	0.0563	
DS X HS X IS	2	0.0677	<1
DS X HS X PS	2	0.1268	2.21
DS X HS X P	6	0.0585	<1
DS X HS X S(P)	40	0.0781	
DS X IS X PS	2	0.0122	<1
DS X IS X P	6	0.2604	<1
DS X IS X S(P)	40	0.0365	
DS X PS X P	6	0.0446	<1
DS X PS X S(P)	40	0.0698	
HS X IS X PS	1	0.0069	<1
HS X IS X P	3	0.1435	1.51
HS X IS X S(P)	20	0.0951	
HS X PS X P	3	0.1435	3.28*
HS X PS X S(P)	20	0.0438	
IS X PS X P	3	0.0046	<1
IS X PS X S(P)	20	0.0229	
DS X HS X IS X PS	2	0.1059	1.85
DS X HS X IS X P	6	0.0677	<1
DS X HS X IS X S(P)	40	0.0920	
DS X HS X PS X P	6	0.0203	<1
DS X HS X PS X S(P)	40	0.0573	
DS X IS X PS X P	6	0.0307	<1

Table 4 (continued)

Analysis of Variance Summary Table for Sensor Error Data

Source	df	Mean Square	F
DS X IS X PS X S(P)	40	0.0698	
HS X IS X PS X P	3	0.1505	<1
HS X IS X PS X S(P)	20	0.0771	
DS X HS X IS X PS X P	6	0.0828	<1
DS X HS X IS X PS X S(P)	40	0.0573	

* $p < .05$

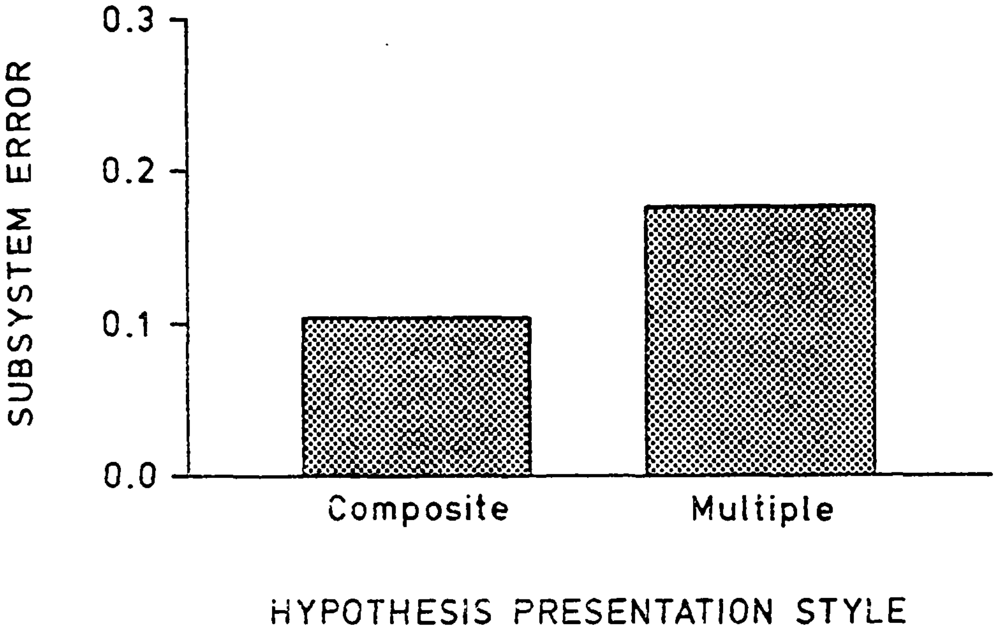


Figure 12. Subsystem Error for Hypothesis Presentation Style

Subjects made fewer errors in the picture-composite condition than in the symbol- (0.156) and text-composite (0.125) conditions, while they committed a greater number of errors in the picture-multiple condition than in the symbol- (0.156) and text-multiple (0.125) conditions. No differences in errors committed were found when symbol-composite and text-composite conditions and symbol-multiple and text-multiple conditions were compared.

Insert Figure 13 about here

Hypothesis Presentation Style by Information Presentation Style by Parameter Presentation Style Effects

A significant hypothesis presentation method by information density by parameter presentation style interaction was obtained for subsystem error. These data are illustrated in Figure 14. A Tukey (hsd) analysis of the means revealed that there was no difference in number of errors subject's made when parameter information was presented in the composite hypothesis format. However, when system parameter information was presented in the multiple hypothesis format, significantly more errors were made in the bargarph/all relevant data (0.250) condition then were made in the numerical/all relevant data (0.111) or composite hypothesis/all relevant (0.056) conditions. A greater number of errors were committed in the multiple hypotheses/all relevant data condition (0.250) then in the composite hypothesis/all relevant condition (0.056). It is clear from the information presented in Figure 14 that the

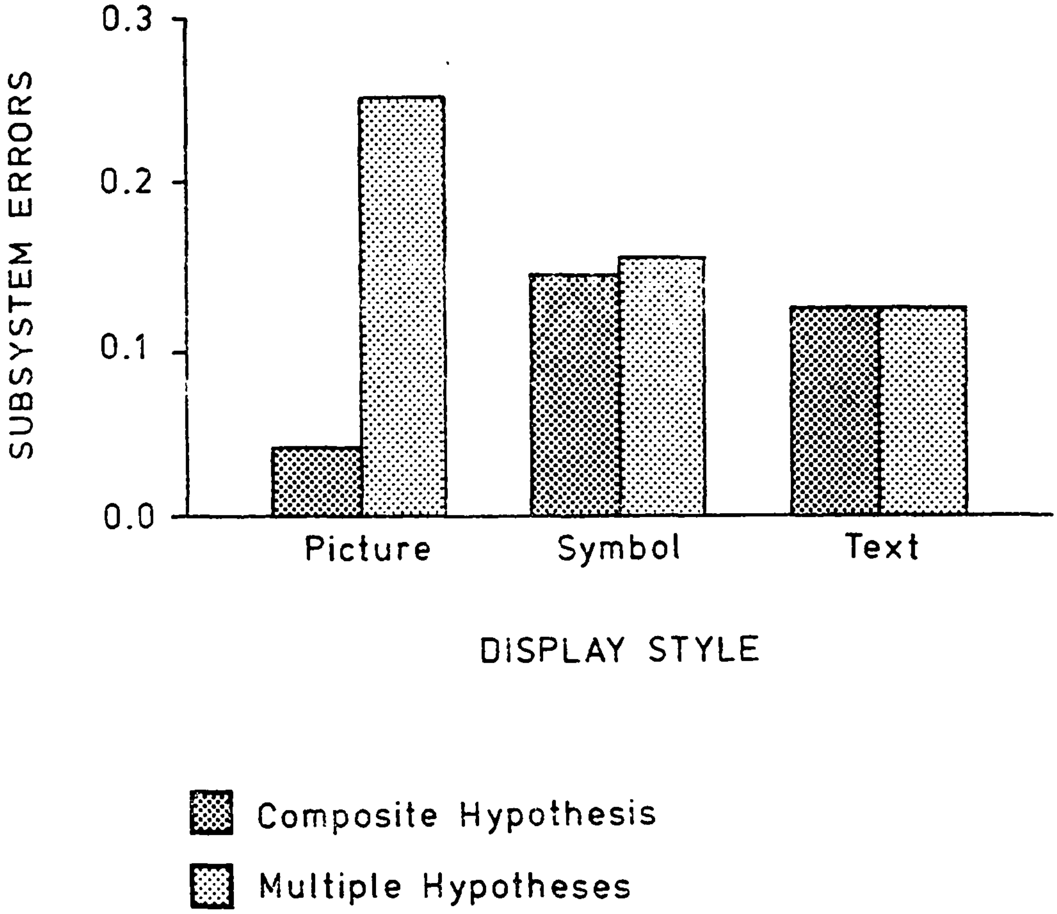


Figure 13. System Errors for Display Style by Hypothesis Presentation Interaction

interaction was the result of the large number of errors committed in the multiple hypothesis format when all the relevant parameters where presented in bargraph form.

Insert Figure 14 about here

Hypothesis Presentation Style by Parameter Presentation Style by Pilot Type Effects

A significant parameter presentation style by hypothesis presentation style by pilot type interaction was observed for sensor errors. These data are represented in Figure 15. No differences were found in the post hoc examination of the means and it is clear from the data presented that there is no evident pattern.

Insert Figure 15 about here

Subjective Data

Upon completing the study subjects were asked to indicate which of the display styles, hypotheses and parameter presentations, and amount of information they preferred based on their experiences with them in the study. In addition subjects were also asked to indicate their reasons for their preferences and to make any recommendation for changes in the displays that might improve pilot performance when using the

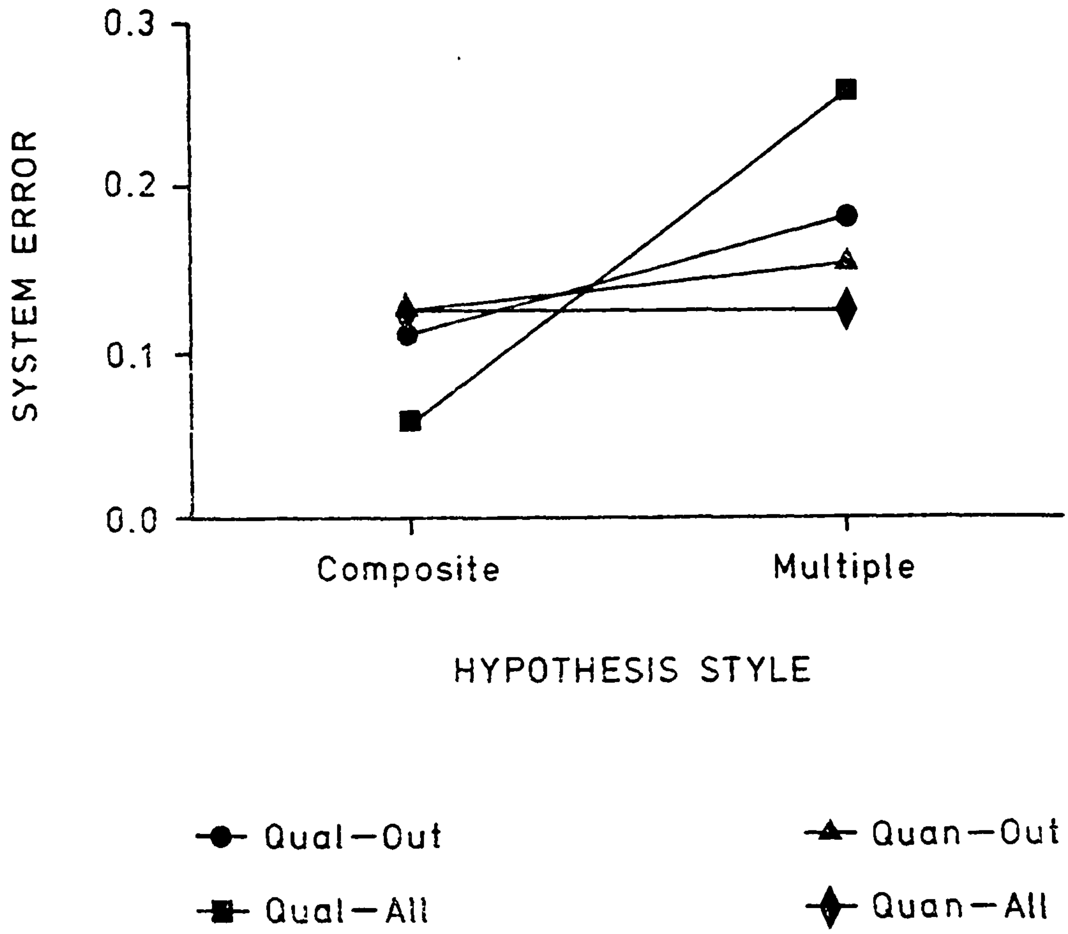


Figure 14. Hypothesis Style, Information Style, and Parameter Style Interaction

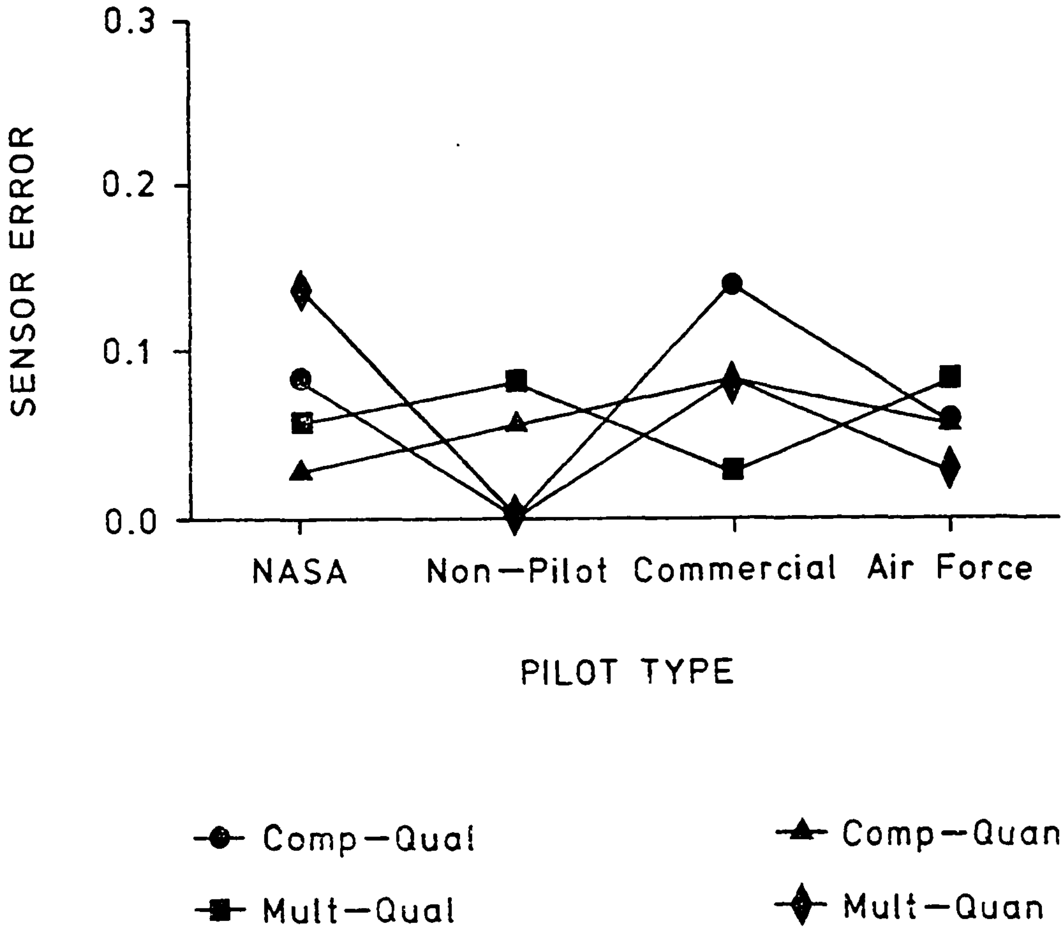


Figure 15. Hypothesis Presentation, Parameter Style, and Pilot Interaction

displays. A copy of the written questionnaire can be found in Appendix E.

Subject Preferences

Subject preference information is presented in Table 5 (Note: responses are presented separately for pilots and non-pilots). Overall subjects preferred the picture-based displays over text- and symbol-based displays. Ease of identification and recognition were the most frequent reasons given by subjects who preferred the picture-based displays, while quick identification and little chance for error were the most frequently given by subjects were preferred the symbol- and text-based displays, respectively. Subjects least preferred the symbol- and text-based, with subjects indicating that the subsystems that corresponded to the symbols were too hard to remember and that the text-based displays took too long to read for a "quick reference."

Insert Table 5 about here

Subjects recommended that the picture used to represent the hydraulic system in the picture-based displays be changed, but were unable to suggest an alternative. Subjects also suggested that the picture used to represent a leak in the picture-based displays be modified to rely less on color to distinguish the different subsystems and more on the picture itself.

Table 5

Subject Preference Data

	Display Style			
	Pilots		Non-Pilots	
	Most Preferred	Least Preferred	Most Preferred	Least Preferred
Picture	10	3	3	0
Symbol	3	8	3	1
Text	5	7	0	5

Hypothesis Presentation	Pilots	Non-Pilots
Composite Hypothesis	9	5
Multiple Hypotheses	9	1

Information Quantity Style	Pilots	Non-Pilots
Out-of-Tolerance Only	11	4
All Relevant	7	2

Parameter Presentation	Pilots	Non-Pilots
Bargraphs	15	5
Numbers	3	1

There was no difference in subjects' preferences for the multiple hypotheses and composite hypothesis displays. Subjects preferring the composite displays indicated that they liked the simplicity and ease of use of these displays, while subjects who preferred the multiple displays indicated that they liked the fact that these displays included more information. When asked why they did not like the non-preferred displays subjects reported that they did not dislike the non-preferred display, but simply were more comfortable with one type of display. A few subjects felt that both types of displays should be included in the system with the pilots able to go from the composite to multiple displays with the push of a button, allowing the pilots to determine the amount of information they would like to have based on current flight conditions.

When asked which of the two styles of parameter information presentation they preferred, subjects indicated a 11 to 7 preference for the presentation of out-of-tolerance information only over the presentation of surplus information. Subject's preference for the out-of-tolerance only style was based upon reports that they felt decisions could be made quicker, and that information about the normal parameters was unnecessary for decision making. Subjects who preferred the surplus information displays reported that they liked this style because decisions are made based upon the status of the entire system, with reassurance that other parameters were normal. As with the hypotheses presentations, a few subjects recommended a two-step procedure be used in the presentation of parameter information. The first step would show

only the out-of-tolerance parameters and, with the press of a button, the pilot could display all of the parameter information.

When asked which of the two parameter presentations they preferred, bargraphs and numerical values, subjects preferred the bargraphs 16 to 2 over the numerical values. Quick recognition and ease of interpretation were the reasons subjects gave for this preference. The need to commit out-of-tolerance values to memory and the lack of need for numbers were cited as the reasons why subjects did not like the numerical parameter presentations.

A modification that was suggested by a few of the pilots was the incorporation of some numerical values into the bargraphs in a way similar to what is being done with the primary flight displays in B767's and A310's where numerical values are combined with bargraphs in the flight displays. A suggestion consistent with this thinking was to include numbers in the displays when the parameter enters the out-of-tolerance range.

When subjects were asked to identify displays where they felt overwhelmed by the information presented, ten subjects reported such experiences. This occurred on displays that presented both multiple hypotheses and surplus parameter information as well as displays that were composed of text and symbols.

Discussion

The rapid development and implementation of computer technology in the area of intelligent machines has surpassed our knowledge of the ramifications of such implementations. Research is needed in the area of human-computer interaction that addresses the specific issues relevant to the implementation of expert system technology into the existing workplace. The current research project was an investigation of different display issues as they relate to the introduction of an expert fault diagnostic system into the commercial aircraft platform. The basic findings suggest that fault explanation hypotheses should be presented in their composite form using a picture-based display format, with only the out-of-tolerance parameter data presented in order to maximize pilot response time and minimize pilot error.

Display Style

It was proposed that of the three styles of displays (picture-, symbol-, and text-based) subjects' best performance would result from their use of the picture-based displays and their worst performance would occur when the symbol-based displays were used. While these results were not directly substantiated by the analysis and examination of subjects' response times and system error data for Display Style, support was provided from the Display Style by Hypothesis Presentation Style interaction. Response times were found to be faster for the picture-composite and the text-composite displays than for the symbol-

composite displays; however, no differences were found between subjects' response times to the picture-composite and text-composite displays. The lack of difference in speed of response is not surprising for these two conditions. In the text-based displays subjects simply had to read the name of the abnormal subsystem(s) from the display while for the picture-based displays subjects had to identify the picture-element representing the abnormal subsystem(s) which contained some text as well as intrinsic meaning. It is expected that had the picture-based/composite displays not contained the text information, response times for the text-based displays would have been faster. This is because subjects did not have enough time to obtain adequate experience with the picture elements used in the study due to the nature of the experimental conditions.

It is expected that with overlearning of the picture elements, subjects' response times will improve appreciably, while the response times for text-based displays would not be expected to improve since subjects would be expected to be already at a state of overlearning. There is experimental evidence that provides additional support for this conclusion since there were no differences in the response times for the different display styles in the multiple hypotheses condition when the picture-based displays did not contain any textual information. One would have expected response times to be faster for the text-based displays in this situation also. However, since there were no differences across the different display styles, this suggests that further training and experience with the picture- and symbol-based

displays would result in quicker response times for them and not for the text-based displays, since overlearning should have already occurred.

In addition to further training, optimization of the picture elements should result in additional improvement in subjects' response times; whereas, such improvements would not be expected in subsequent modifications of the text-based displays, since the key elements, e.g., words, would not be changed.

It was predicted that subjects' response times would be slowest using the symbol-based displays. This hypothesis was supported by the Display Style by Hypothesis Presentation Style results that indicated a longer response time for symbol-based displays using the composite-hypothesis presentation style. This difference may be attributed to the inclusion of text in the picture-composite symbols which allowed subjects to rely less on their memories and more on the text included in the display.

Examining the individual cases in the case by display style by hypothesis presentation style interaction, it was interesting to note that in the simpler cases (A, B, C, and D), those involving only one subsystem and one or two hypotheses, subjects' response times did not significantly differ across the display style or hypothesis presentation style. However, when the amount of information presented in the displays increased, the number of subsystems increased to two or three and/or the number of hypothesis increased, response time differences

occurred. In Cases E, F, G, and H, subjects' response times were quickest in the picture-composite and text-composite conditions and slowest in the symbol-multiple conditions, providing further support for the experimental hypothesis.

The commission of errors was less frequent in the picture-composite condition then in any of the other conditions; however, error frequency was greatest in the picture-multiple condition. When the types of errors subjects made in the picture-based displays were looked at in greater detail it was apparent that the use of color as the only indication of the abnormally operating subsystem when a leak occurred, was difficult for the subjects. Questioning of the subjects revealed that the colors used to represent the fluids of the different subsystems did not correspond to the subjects' expectations of what those colors should be. Since no consistent color scheme was proposed and it is expected that the Federal Aeronautics Administration would not approve a class of pictures differentiated solely by color, a set of display elements need to be developed that eliminate the sole reliance on color in these displays and use the pictures themselves to differentiate among the various subsystems.

Subjects preferred the picture-based displays 2 to 1 over the text-based displays, and the text-based and symbol-based displays were the least preferred almost equally, 7 and 8 respectively. Subjects reported preferring the picture-based displays because they took less time to interpret and were easiest to identify, while they felt that the text-

based displays took too long to read and the symbol-based displays required too much interpretation. Since there is a clear preference toward the picture-based displays and reason to expect improvement with practice and display optimization that is not expected with the text-based displays, the a picture-based display format would appear to be the logical choice for display style. These findings are supportive of the work of Summers and Erickson (1984) and Way et. al. (1984) who have been attempting to implement picture-based displays in the cockpit of fighter aircraft.

When further research to generate new picture elements is undertaken it will be important to consider how they relate to the other displays that will be included in the cockpit. While this is seen as a limitation to the current research, it was undertaken anyhow without any considering of other concurrent displays since they are as yet unavailable.

Hypothesis Presentation Style

The comparison of response times to different hypothesis presentation styles was born from the computational approach taken by Faultfinder which generates hypotheses to explain detected out-of-tolerance conditions (i.e. faults). Two presentation styles were created in order to address this issue. The multiple hypothesis style was based on the assumption that pilots might be more interested in information on the individual hypothesis that Faultfinder was working on at any given time during its diagnostic computations. This set of

displays provides information about components of a subsystem that could be producing the current abnormal conditions. The composite hypothesis presentation style simply identifies the subsystem where a suspected fault has been detected and does not address subsystem components. Clearly the multiple hypotheses displays contain more information, and one would expect their response times to be greater (Goldsmith, 1981). It is not then surprising that the results of this study substantiate this expectation. Response times, overall, were consistently faster in the composite hypothesis displays (4.530 sec) compared to the multiple hypothesis displays (5.370 sec). Where differences occurred in the post hoc tests, the response times for the composite hypothesis displays were consistently faster than those for the multiple hypothesis displays.

The differences found between the two types of hypothesis presentation style were not consistent across the different display styles. The experimental results indicate that although subjects' response times were faster in the composite hypothesis displays for symbol-based display style, they were not significantly faster in all but one of the multiple hypotheses cases (see Figure 5 and 6).

When asked their preference between the two hypothesis presentation styles, pilots were split evenly between the two, but the non-pilots showed a clear preference for the composite displays. Discussion with some of the pilots indicated that they felt a need for the speed associated with the composite displays as well as a need for the component information contained in the multiple hypotheses displays. A

viable solution to this situation, suggested by a few of the pilots, is to include both presentation styles in the interface. A scheme for the application of this suggestion is to present the composite presentation style when a out-of-tolerance condition is first detected and allow the pilot to switch to the multiple hypotheses display by performing some action (i.e., press button, voice command, etc.) when workload permits. Such a scheme would allow for quick detection of faulty conditions during critical phases of flight (takeoffs and landings) through the use of the composite displays, and enable the pilot to access more specific information when workload is at a more optimum level for decision making. This is consistent with the suggestions made by Klein and Calderwood (1986) concerning the optimization of information presentation to operators in time-critical situations and also with the suggestions made by Schmit (1984) and Murphy and Mitchell (1986) to allow pilots to control the level of display detail.

Parameter Presentation Style

The results of the current study showed no overall differences between subjects' response times for the two parameter presentation styles. One style of parameter information presentation represented parameter information as a bargraph, with a needle indication current status. The second type of parameter presentation represented information as a number with background color and arrow heads to indicate current parameter status. This finding appears to be the result of the form of presentation used in the investigation rather than differences between the two types of parameter displays investigated.

The task of the subject was to identify the operating status (normal, abnormal-high, or abnormal-low) of the individual parameters. Even though the appearance of the two parameters differ, the task the subjects had to perform to obtain the operating status of the parameter was very similar. For the bargraph displays, this simply meant locating the position of the white bar and identifying the region of the bar where it was positioned. In the numerical value display, subjects could simply identify the background color of the box (easier to do in the numerical value displays than the bargraph displays) and then look at the arrowhead in the red boxes (more difficult than for the bargraphs) to determine if the parameters were high or low. Thus, the number presented in the box was unimportant to the subject and for the most part was ignored by them. The required response of the subject was essentially the same for the two types of the parameter displays and so the lack of difference in response time. These circumstances would be different in current aircraft cockpits; there pilots have to memorize critical values for the different parameter values based on the different phases of flight. Recall that these conditions can at times be difficult, straining short-term memory, and can result in error.

Pilots preferred the bargraphs to the numerical values five to one. They reported that bargraphs were conducive to recognition "at-a-glance" and were easy to understand and interpret. On the other hand, numbers contained in the numerical displays required memorization and were unimportant. It is recognized, however, that there will be some instances where numerical values have some importance to the pilot, such

as low oil, fuel, and hydraulic quantities. It is recommended, based on the experimental results and subjects' preferences, that bargraphs are the appropriate choice for inclusion in the displays. However, it is suggested that further research is necessary to determine how and when it would be appropriate to include numbers in parameter displays.

Information Presentation Style

It was hypothesized that subjects performance would be better in conditions where only out-of-tolerance information was included in the displays compared to situations where all relevant information was presented. The results of the current investigation support this prediction. Subjects responded more quickly to out-of-tolerance only displays (4.283 sec) than to all relevant displays (5.616 sec). This finding was consistent across the eight test cases (see Figure 4). This hypothesis is further supported by the finding of the case by information presentation style by parameter presentation style interaction. Where differences occurred, performance was faster for the out-of-tolerance only displays.

Pilot preference was also in support of the out-of-tolerance only displays (11 to 7). Pilots preferred this type of display because they felt it took less time to identify the out-of-tolerance parameters and decisions could be made quicker when normal parameter information, unnecessary for the decision, was eliminated from the display. The finding that subjects' responses were quicker in the out-of-tolerance only condition are consistent with basic cognitive research that has

demonstrated that response times are longer when additional information is added to the search space. It is not clear from this investigation what role normal parameter information might play in pilot decision making in more realistic settings. It is therefore suggested that the displays used in Faultfinder be generated based upon the out-of-tolerance only presentation style; however, further research needs to be undertaken to determine when it might be appropriate to display parameters that are approaching abnormality (i.e. in the yellow), a situation not addressed in the present investigation.

Conclusions and Recommendations

The results of the present investigation indicate that future research and development of the interface to Faultfinder should concentrate on picture-based displays. Although subjects' responses to picture- and text-based displays were comparable, it is felt that with optimization of the picture elements and the subsequent overlearning of these elements will result in marked improvement in response times and accuracy. Comparable improvements cannot be expected with the overlearning of text-based displays since subject familiarity with this media should be a stable level. The intrinsic meaning associated with pictures allows for quicker response time compared with text, which requires interpretation to obtain meaning. This result is consistent with previous research comparing graphic and text displays (Sharit, 1985; Stern, 1984; Tullis, 1981).

Further research into optimizing the picture elements should consider how they would interact with the other displays that will also be found in future generation aircraft. Research should avoid the use of geometric figures in the displays because, as it was pointed out by some of the pilots in the study, such figures are used in current navigational displays and so their use in other systems could result in undesirable confusion among the different systems. Research is needed to determine the role that text could play in picture-based displays since the current result included textual clues in the picture-composite display elements.

For the presentation of the hypotheses generated by Faultfinder it is suggested that further investigation be undertaken to develop a two stage process for presenting them to the pilot. The first stage in the process would be based on the composite hypothesis presentation style where only subsystems are identified as faults. The pilot would then switch to a multiple hypothesis display, one that addresses subsystem component failure, when conditions in the cockpit allow for it. The method for switching between the two displays is a topic for subsequent research.

Results on parameter presentation style did not reveal differences between the two styles employed in the study. This was felt to be a product of the methods used and not a true reflection of the issue at hand. Current practice in the airline industry has been towards the use of bargraphs or strictly qualitative displays instead of numerical readouts because of the questioned usefulness of those numbers. The present research was unable to address this issue since the values included in the display held no meaning for the subjects. It has therefore been suggested that further research be undertaken to determine the viability of including numerical values in bargraphs for critical system parameters.

Finally, the current study demonstrated that subjects were able to perform better (quicker response times) when only the out-of-tolerance parameters were included in the displays, compared to the inclusion of all relevant parameters. This research was consistent with previous

research in the area of cognition and made intuitive sense in that displays that included greater degrees of information density had corresponding increases in response times. It is been suggested that the interface to Faultfinder include out-of-tolerance parameters only as long as research into the inclusion of parameters approaching abnormality be undertaken and that some steps be taken to assure that the all relevant information is available to the pilot in other subsystems displays included in the cockpit.

Future research should include the addition of checklists to the displays. The inclusion of a checklist will become critical to the crew when Faultfinder has narrowed the number of potential fault causes to a point where a reasonable response could be addressed through a checklist. Other relevant information, such as the identification of inoperable equipment and information about the nearest suitable airports, may be appropriate for inclusion in the Faultfinder interface and should be addressed by future research in this area.

References

Abbott, K.H., Schutte, P.C., Palmer, M.T., & Ricks, W.R. (1987). Faultfinder: A diagnostic expert system with graceful degradation for onboard aircraft applications. Paper presented at 14th International Symposium on Aircraft Integration Monitoring Systems. Friedrichshafen, West Germany.

Adam, E. (1987). Tactical aircraft displays. Seminar Lecture Notes, Society for Information Display (9.1-9.19) New Orleans, LA.

Benbasat, I., Dexter, A.S., & Todd, P. (1986). An experimental program investigating color-enhanced and graphical information presentation: An integration of the findings. Communications of the ACM, 29, 1094-1105.

Bergeron, H.B. (1981). General aviation single pilot IFR autopilot study. Presented at the 1980 Aircraft Safety and Operating Problems Conference, Hampton, VA.

Bergeron, H.P. & Hinton, D.A. (1983). Aircraft automation: The problem of the pilot interface. Proceedings of the 2nd Symposium on Aviation Psychology, 95-102.

Blomberg, R.D. & Pepler, R.A. (1983). Performance evaluation of electronic flight instruments. Proceedings of the 2nd Symposium on Aviation Psychology, 17-25.

Booher, H.R. (1975). Relative comprehensibility of pictorial information and printed words in proceduralized instructions. Human Factors, 17, 266-277.

Boucek, G.P., Hanson, D.C., Po-Chedley, D.A., Berson, B.L., Leffler, M.F.,and Hendrickson, J.F. (1981). Aircraft alerting systems standardization study. Digital Avionics Systems Conference, Collection of Technical Papers, 119-128.

Buchanan, B.G. & Feigenbaum, E.A. (1978). Dendral and Meta-Dendral: Their application dimension. Artificial Intelligence, 11, 5-24.

Calhoun, G.L. & Herron, S. (1981). Computer generated cockpit engine displays. Proceedings of the Human Factors Society - 25th Annual Meeting, 127-131.

Chorley, R.A. (1981). Electronic flight deck displays for transport aircraft. Presented at the 4th AIAA/IEEE Digital Avionics Systems Conference - A Collection of Technical Papers (pp. 237-246). St. Louis, MI.

Christ, R.E. (1977). Four years of color research for visual displays. Proceedings of the Human Factors Society - 21th Annual Meeting, 319-321.

Chu, Y. & Rouse, W.B. (1979). Adaptive allocation of decision making responsibility between human and computer in multitask situations. IEEE Transactions on Systems, Man, and Cybernetics, SMC-9, 769-778.

Dickson, G.W., Desanctis, G., & McBride, D.J. (1986). Understanding the effectiveness of computer graphics for decision support: A cumulative experimental approach. Communications of the ACM, 29, 40-47.

Eberts, R.E., Nof, S.Y., Zimolong, B., & Salvendy, G.(1984). Dynamic process control: Cognitive requirements and expert systems. In G. Salvendy's (Ed.), Human-Computer Interaction, (pp. 215-228). Elsevier Science Publishers: Amsterdam, Netherlands.

Graham, & Broomhead, (1975). Cockpit Switching Study, Phase 3, Technology and Interface Requirements (Report AD A015 301). Seattle, WA: Boeing Aircraft Co.

Hanson, D.C., Boucek, G.P., Smith, W.D., Chikos, S.F.,Hendrickson, J.F., Howison, W.W., & Breson, B.L. (1982). Flight operations safety monitoring effects on the crew alerting system. Proceedings of the Human Factors Society - 26th Annual Meeting, 783-787.

Hart, S.G. & Sheridan, T.B. (1984). Pilot workload, performance, and aircraft automation. AGARD Conference Proceedings No. 371, Human Factors Considerations in High Performance Aircraft (pp. 18.1-18.12). NATO, Neuilly Sur Seine, France.

Hoecker, D.G. (1982). Some performance correlates of technical illustration level. Proceedings of the Human Factors Society - 26th Annual Meeting, 88-92.

Hussey, D.W. (1980). Colour displays: Their availability, performance, and application to improved crew efficiency. In AGARD Conference Proceedings No. 267, High Speed Low-level Flight: Aircrew Factors. NATO Neuilly sur Seine, France.

Jauer, R.A. & Quinn, T.J. (1982). Pictorial Formats (Report AFWAL-TR-81-3156). Flight Dynamics Laboratory, Wright-Patterson Air Force Base, OH.

Kellogg, R.S., Kennedy, R.S., & Woodruff, R.R. (1981). A comparison of color versus black and white visual display as indicated by bombing performance in the 2B35 TA-4J Flight Simulator. Proceedings of the Human Factors Society - 25th Annual Meeting, 233-234.

Klein, G. & Calderwood, R. (1986). Human factors considerations for expert systems. Proceedings of IEEE Conference on Systems, Man, and Cybernetics, 921-925.

Krebs, M.J., Wolf, J.D., & Sandvig, S.H. (1978). Color Display Design Guide. (Office of Naval Research Report Number ONR-CR-213-136-2F). San Diego, CA: Office of Naval Research.

Kopala, C.J. (1979). The use of color-coded symbols in highly dense situation displays. Proceedings of the Human Factors Society - 23rd Annual Meeting, 397-401.

Mas, G.E., Erickson, J.B., & Jordan, D. (1979). Engine Monitoring and Display System (EMADS) (Douglas Paper 6851). Long Beach, CA: Douglas Aircraft Company.

Moss, R.W. & Reising, J.M. (1984). Cockpit automation: Design and implementation issues. Proceedings of the 3rd Aerospace Engineering and Technology Conference, 41-45.

Moss, R.W., Reising, J.M., & Hudson, N.R. (1984). Automation in the cockpit: Who's in charge? Proceedings of the 3rd Aerospace and Engineering Technology Conference, 25-29.

Murphy, E.D. & Mitchell, C.M. (1986). Cognitive attributes: Implications for display design in supervisory control systems. International Journal of Man-Machine Studies, 25, 411-438.

Patterson, M.J., Engleman, W.J., Najjar, L.W., & Corso, G.M. (1984). The effects of redundancy, target density, and practice on a complex visual search task. Proceedings of the Human Factors Society - 28th Annual Conference, 299-302.

Peterson, R.J., Banks, W.W., & Gertman, D.I. (1981). Performance-based evaluation of graphic displays for nuclear power plant control rooms. Proceeding of SIG-CHI, 182-189.

Rasmussen, J. (1983). Skills, rules, knowledge; Signals, signs, and other distinctions in human performance models. IEEE Transactions on Systems, Man, and Cybernetics, SMC-13, 257-266.

Reising, J.M. & Artez, A.J. (1983). Color coding in fighter cockpits: It isn't black and white. Proceedings of the 2nd Symposium on Aviation Psychology, 1-7.

Reising, J.M. & Calhoun, G.L. (1982). Color display formats in the cockpit: Who needs them? Proceedings of Human Factors Society - 26th Annual Meeting, 446-449.

Robinson, C.P. & Eberts, R.E. (1987). Comparison of speech and pictorial displays in a cockpit environment. Human Factors, 29, 31-44.

Rockwell, T.H., Giffin, W.C., & Romer, D.J. (1983). Combining destination diversion decisions and critical inflight event diagnosis in computer aided testing of pilots. Proceedings of the 2nd Symposium on Aviation Psychology, 343-351.

Rouse, S.H., Rouse, W.B., & Hammer, J.H. (1982). Design and evaluation of an onboard computer-based information system aircraft. IEEE Transactions on Systems, Man, and Cybernetics, SMC-12, 451-463.

Rouse, W.B. (1981). Pilot Interaction With Automated Airbourne Decision Making Systems (Report T-106). Illinois University.

Rouse, W.B., Geddes, N.D., & Curry, R.E. (1987). An architecture for intelligent interfaces: Outline of an approach to supporting operators of complex systems. Human-Computer Interaction, 3.

Schmit, V.P. (1984). Changing system/display concepts and their impact on aircrew performance. AGARD Conference Proceedings No. 371, Human Factors Considerations in High Performance Aircraft (pp. 13.1-13.7). NATO, Neuilly Sur Seine, France.

Schneider, E.T. & Enevoldson, E.K. (1984). Modernizing engine displays. Proceedings of the 6th Advanced Aircew Display Symposium (pp. 96-103). Patuxent River, MD.

Schudy, R., Corker, K., Baron, S. (1987). Intelligent interfaces to aircraft systems. Paper submitted to the AIAA Conference on Computers in Aerospace.

Sexton, G.A. (1983). Crew Systems and Flight Station Concepts for a 1995 Transport Aircraft (NASA Contractor Report 166068). Lockheed-Georgia Company, Marietta, GA.

Sharit, J. (1985). Supervisory control of a flexible manufacturing system. Human Factors, 27, 47-59.

Shortliffe, E.H. (1976). Computer-Based Medical Consultation: Mycin. New York: American-Elsevier.

Stern, K.R. (1984). An evaluation of written, graphics, and voice messages in proceduralized instruction. Proceedings of the Human Factors Society - 28th Annual Meeting, 314-318.

Summers, L.G. & Erickson, J.B. (1984). System Status Display Information (NASA Contractor Report 172347). Hampton, VA: National Aeronautics and Space Administration.

Tullis, T.S. (1981). An evaluation of alphanumeric, graphic, and color information displays. Human Factors, 23, 541-550.

Veitengruber, J.E. (1977). Design criteria for aircraft warning, cuation, and advisory alerting systems. Proceedings AIAA Aircraft Systems and Technology Meeting, 1-8. Seattle, WA.

Wanner, J.C. (1982). Human factors and flight safety. In B. Laschka and R. Staufenbiel (Eds.), Proceedings of the 13th Congress of the International Council of the Aeronautical Sciences (pp. 83-96). Seattle, WA.

Way, T.C., Hornsby, M.E., Gilmour, J.D., Edwards, R.E., & Hobbs, R.E. (1984). Pictorial Format Display Evaluation (Report AFWAL-TR-84-3036). Flight Dynamics Laboratory, Wright-Paterson Air Force Base, OH.

Weiner, E.f. (1984). Flight deck automation. Proceedings of the Aerospace Engineering and Technology Conference, 3-8.

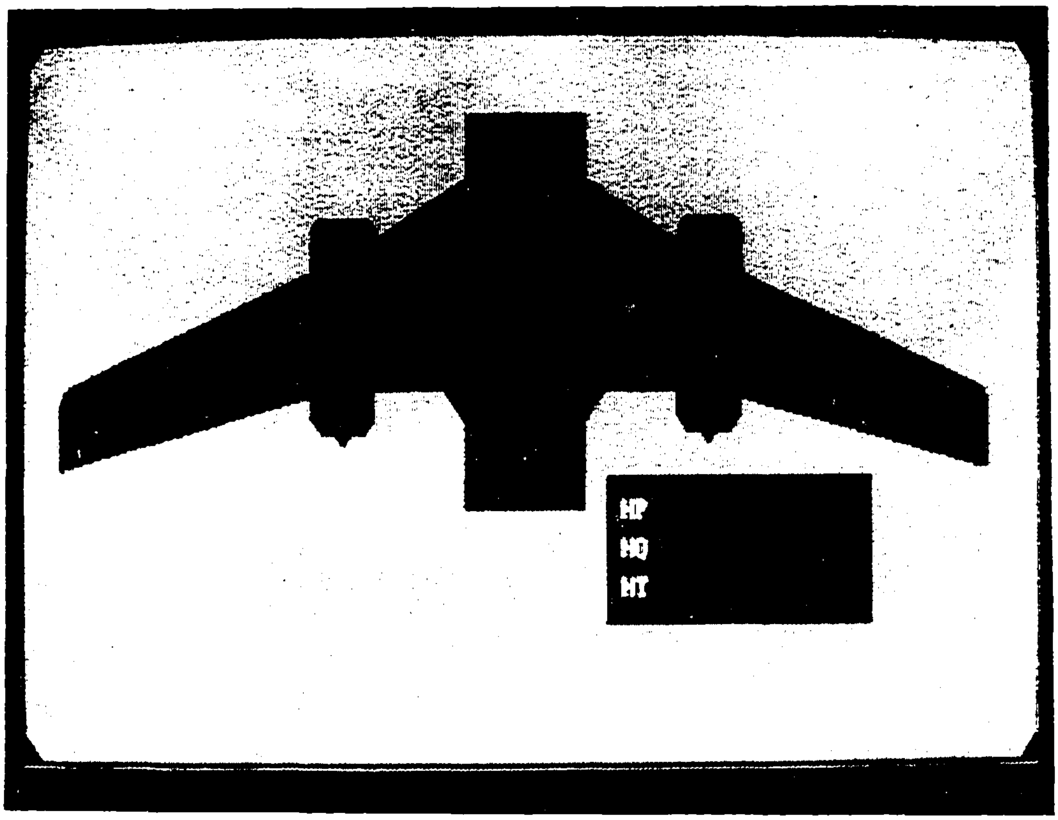
Appendix A

Experimental Displays

Displays for Case A

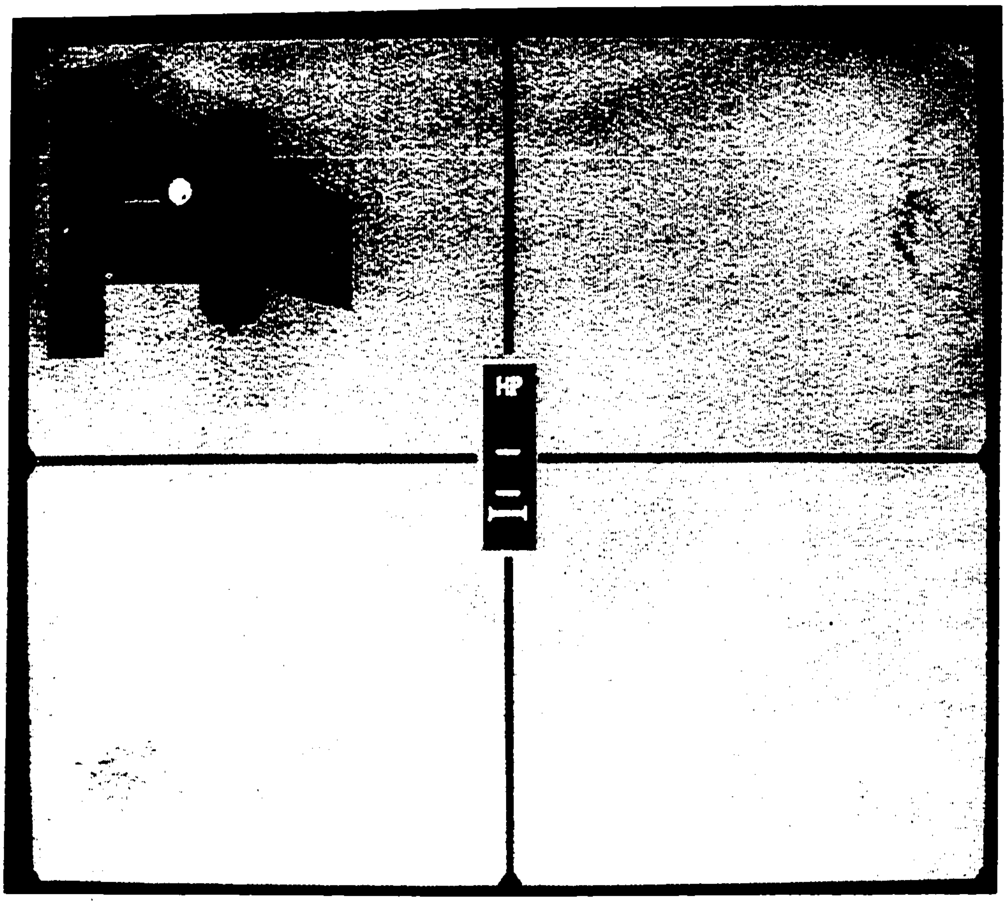
Right Hydraulic Pump Failure

(Appendix Continued)



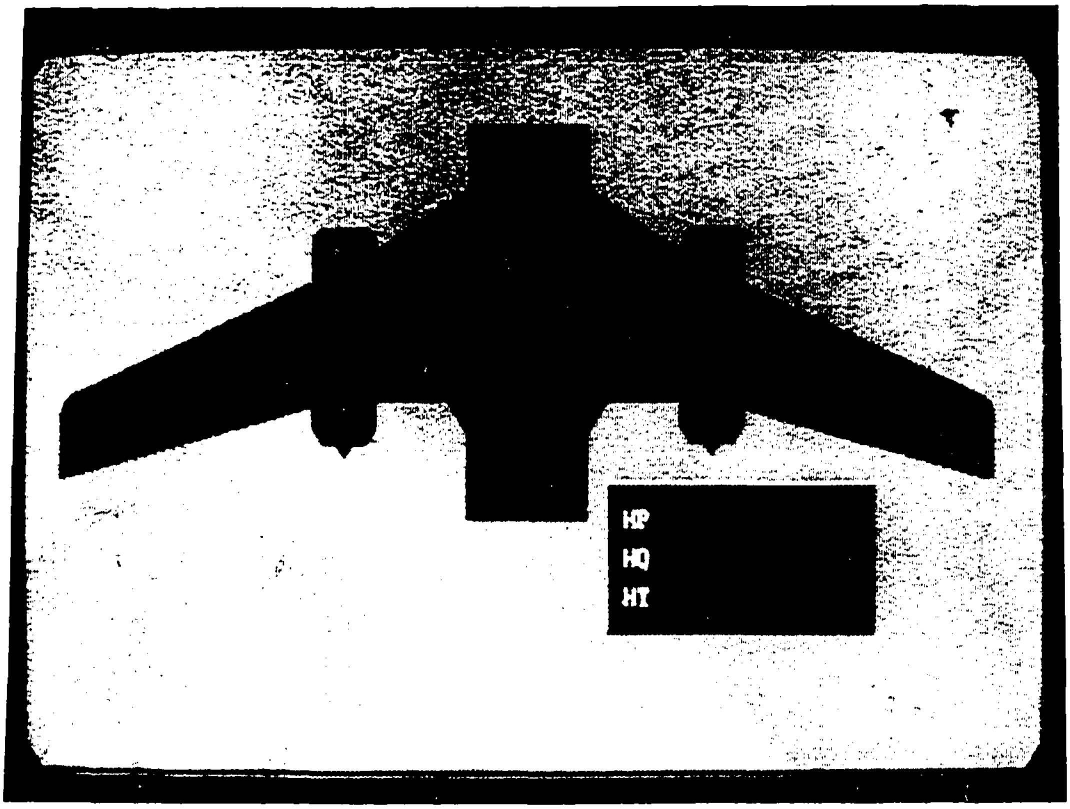
Picture - Composite Hypothesis

(Appendix Continued)



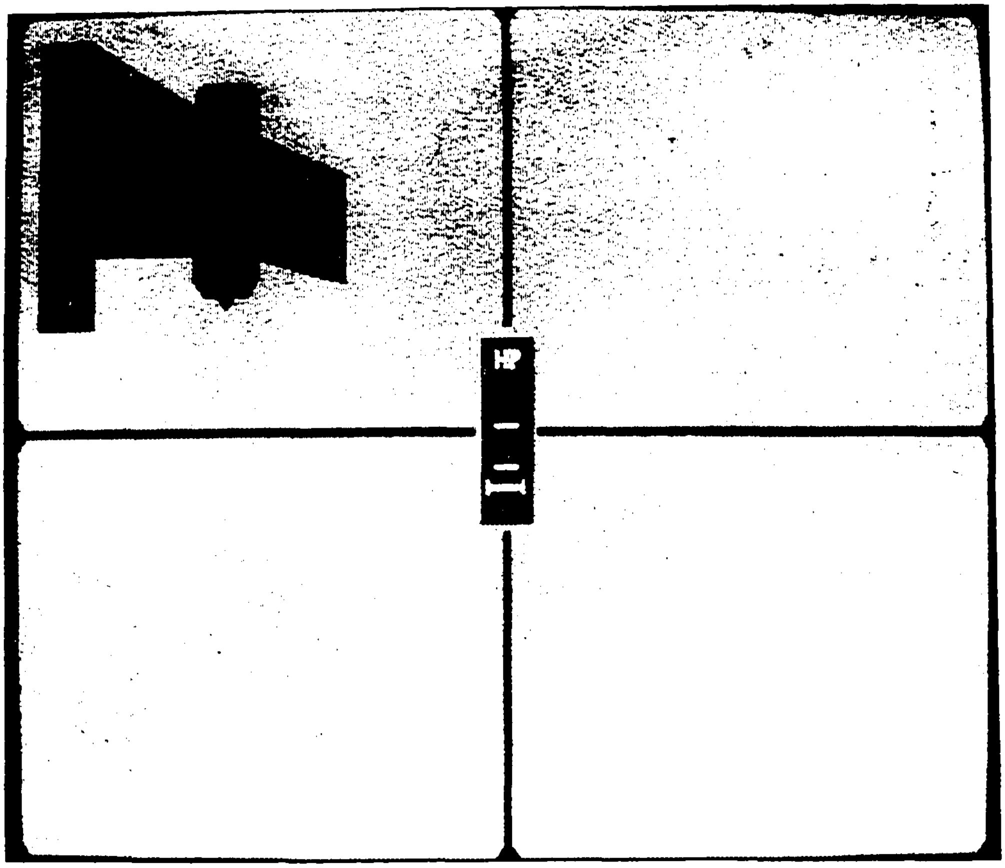
Picture - Multiple Hypotheses

(Appendix Continued)



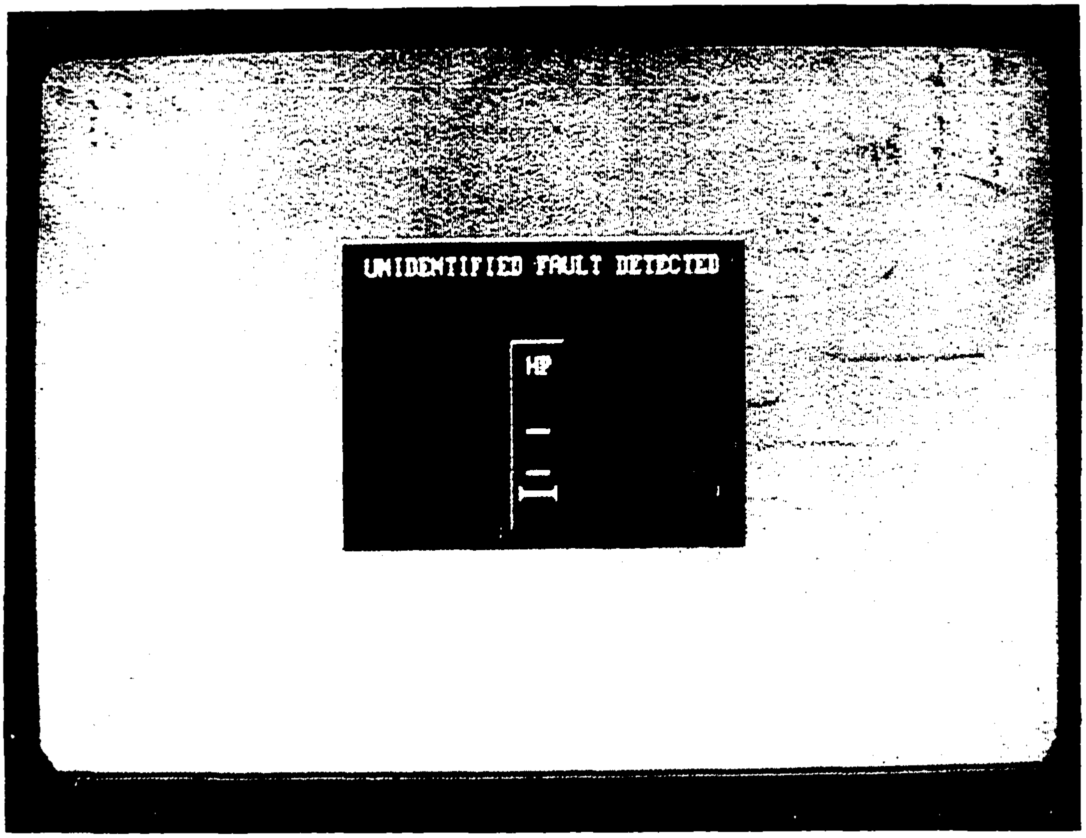
Symbol - Composite Hypothesis

(Appendix Continued)



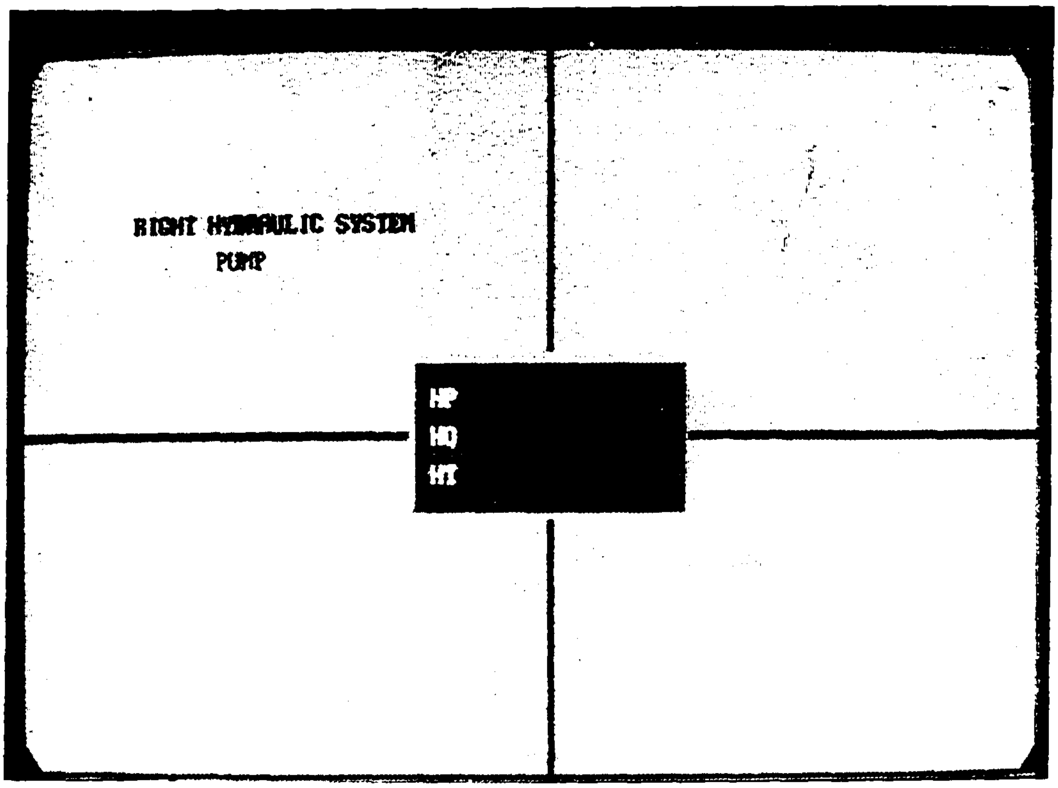
Symbol - Multiple Hypotheses

(Appendix Continued)



Text - Composite Hypothesis

(Appendix Continued)



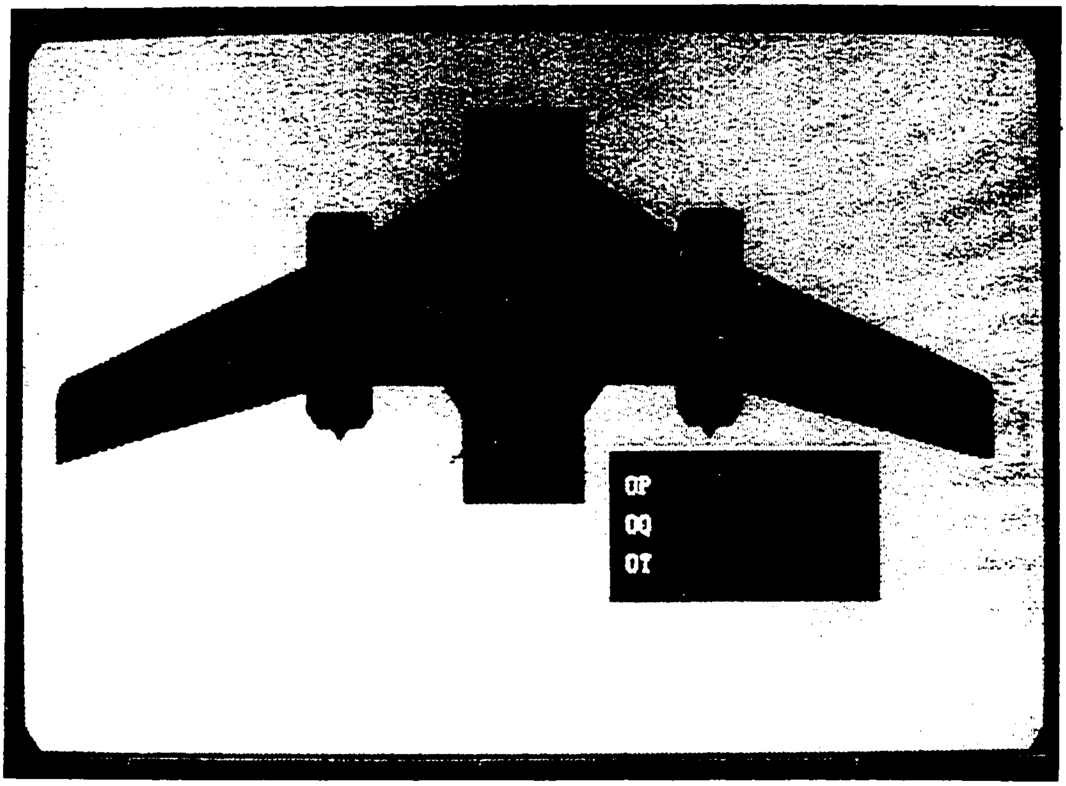
Text - Multiple Hypotheses

(Appendix Continued)

Displays for Case B

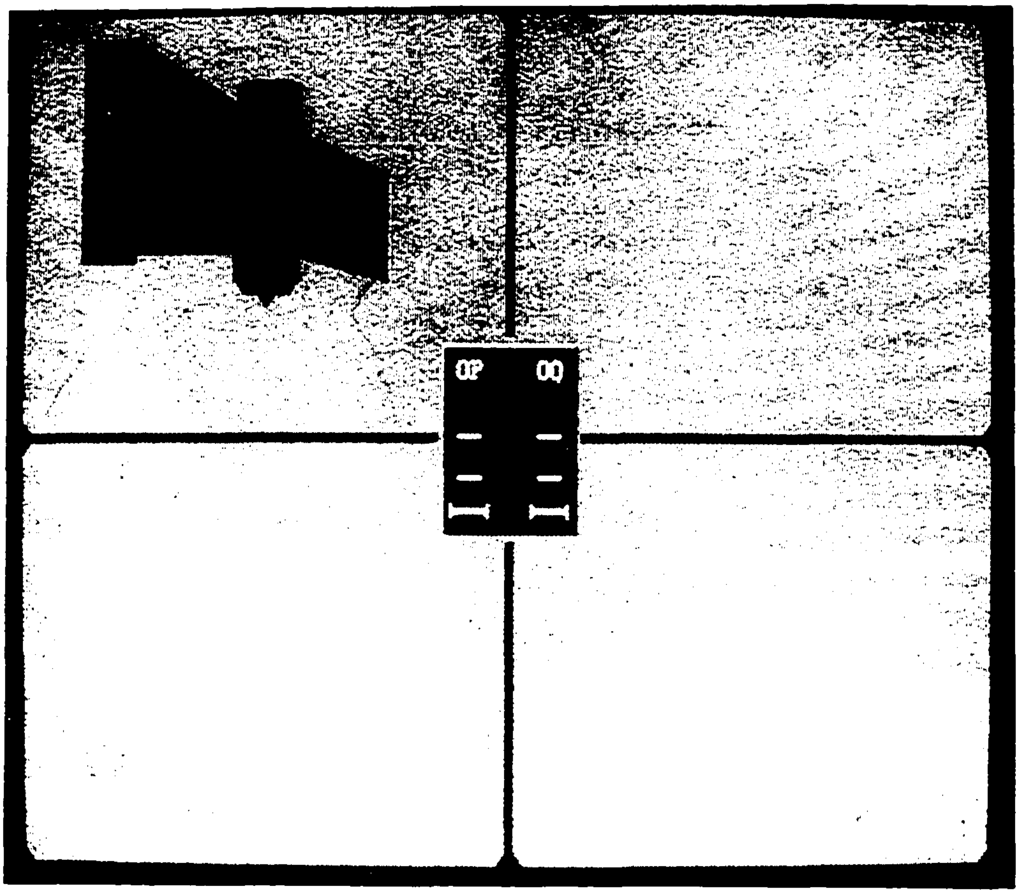
Left Oil Leak

(Appendix Continued)



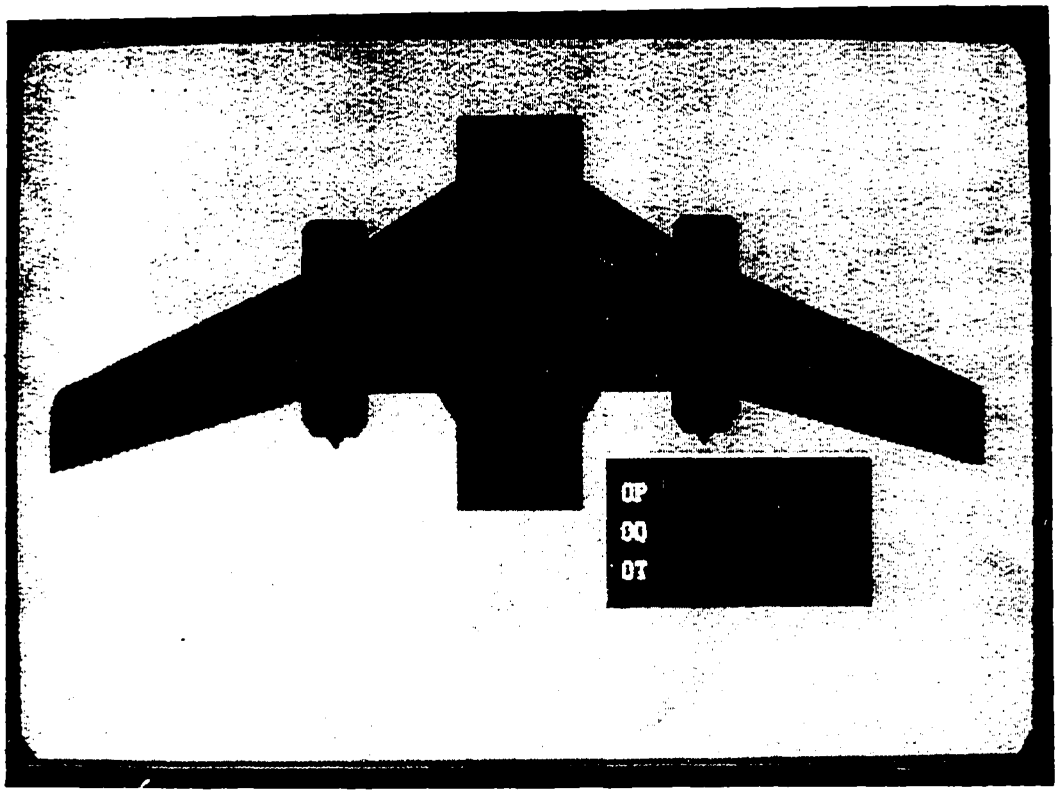
Picture - Composite Hypothesis

(Appendix Continued)



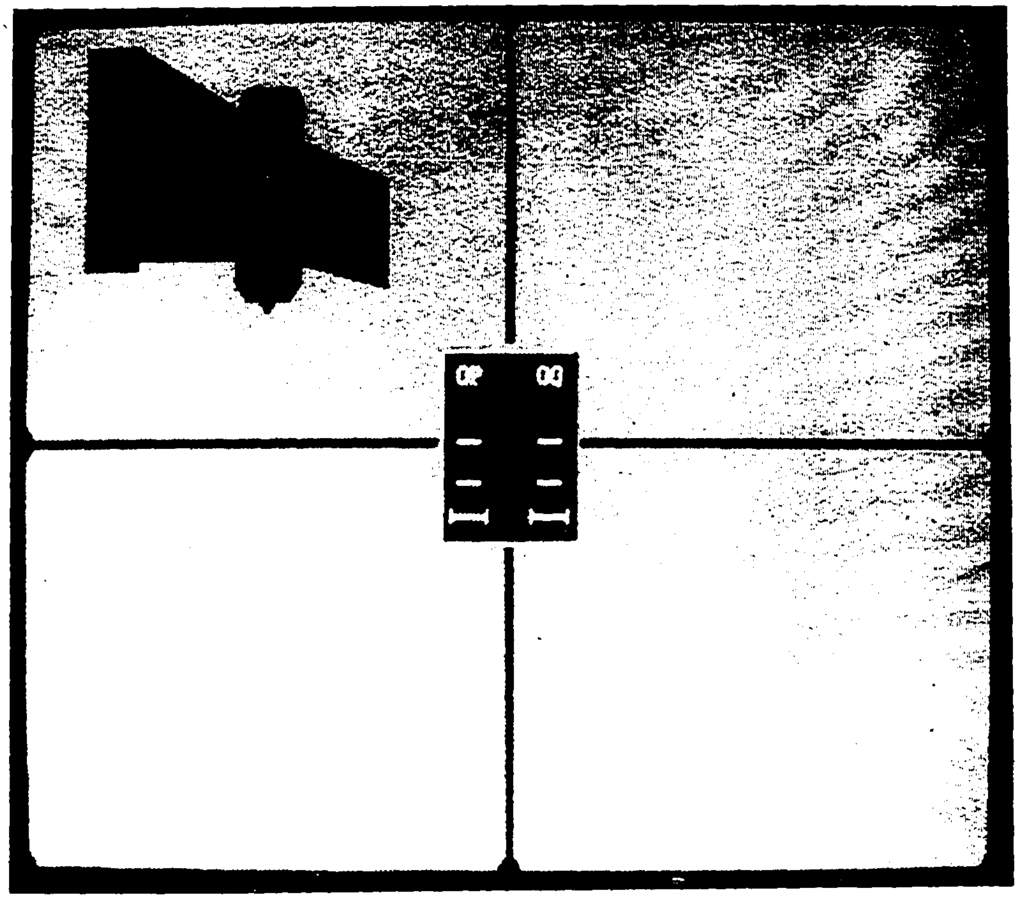
Picture - Multiple Hypotheses

(Appendix Continued)



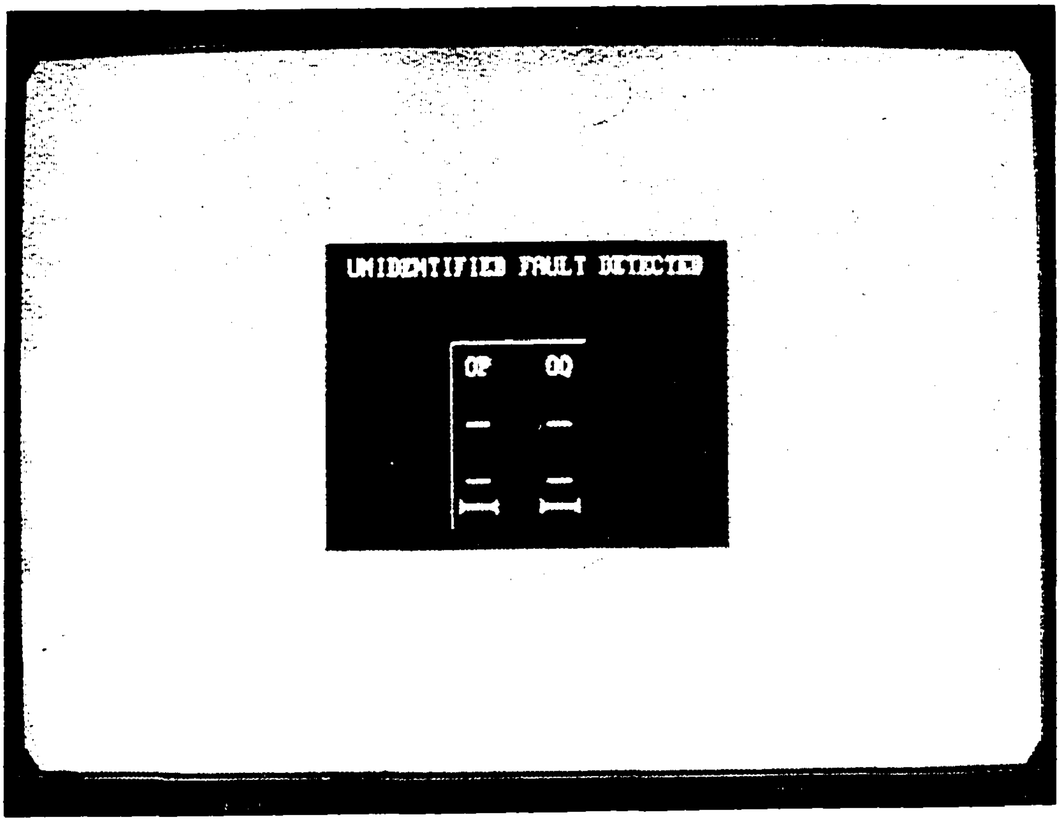
Symbol - Composite Hypothesis

(Appendix Continued)



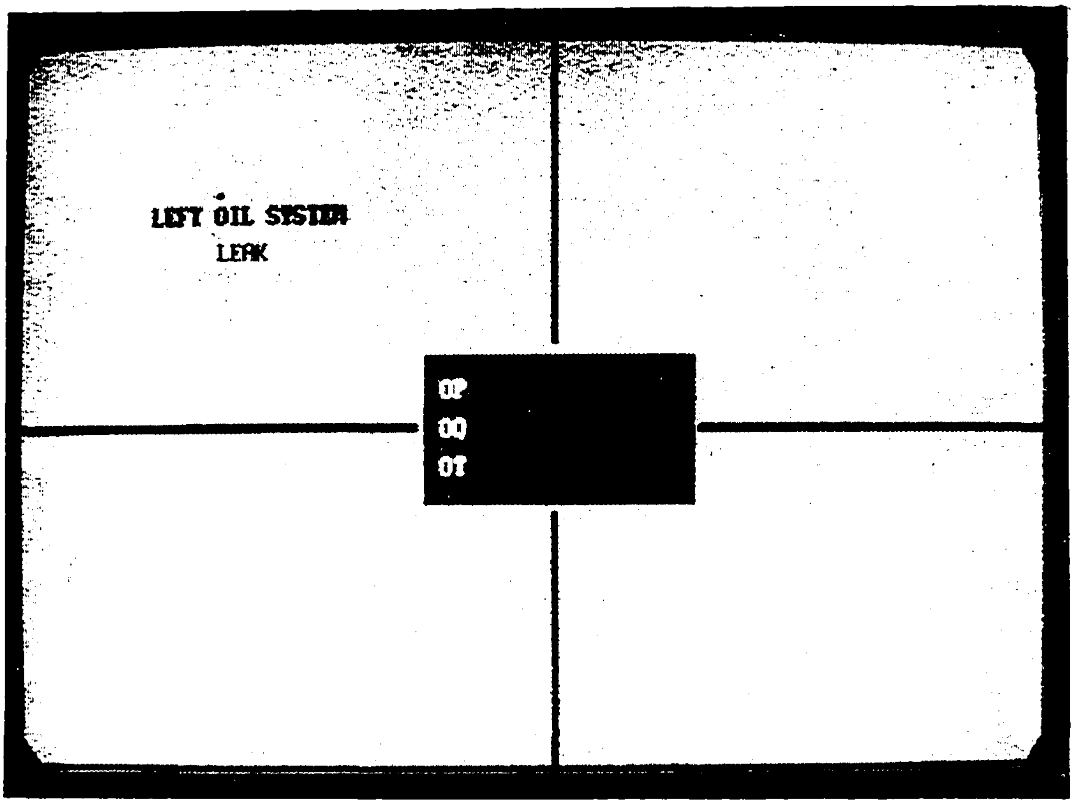
Symbol - Multiple Hypotheses

(Appendix Continued)



Text - Composite Hypothesis

(Appendix Continued)



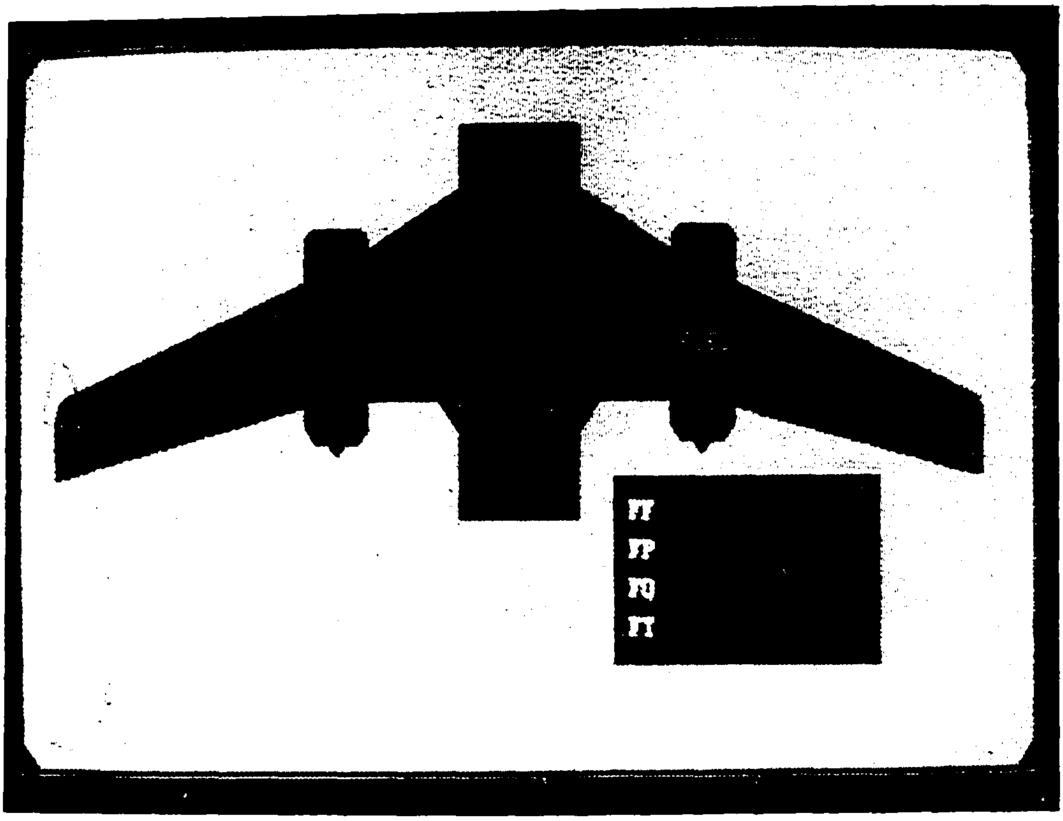
Text - Multiple Hypotheses

(Appendix Continued)

Displays for Case C

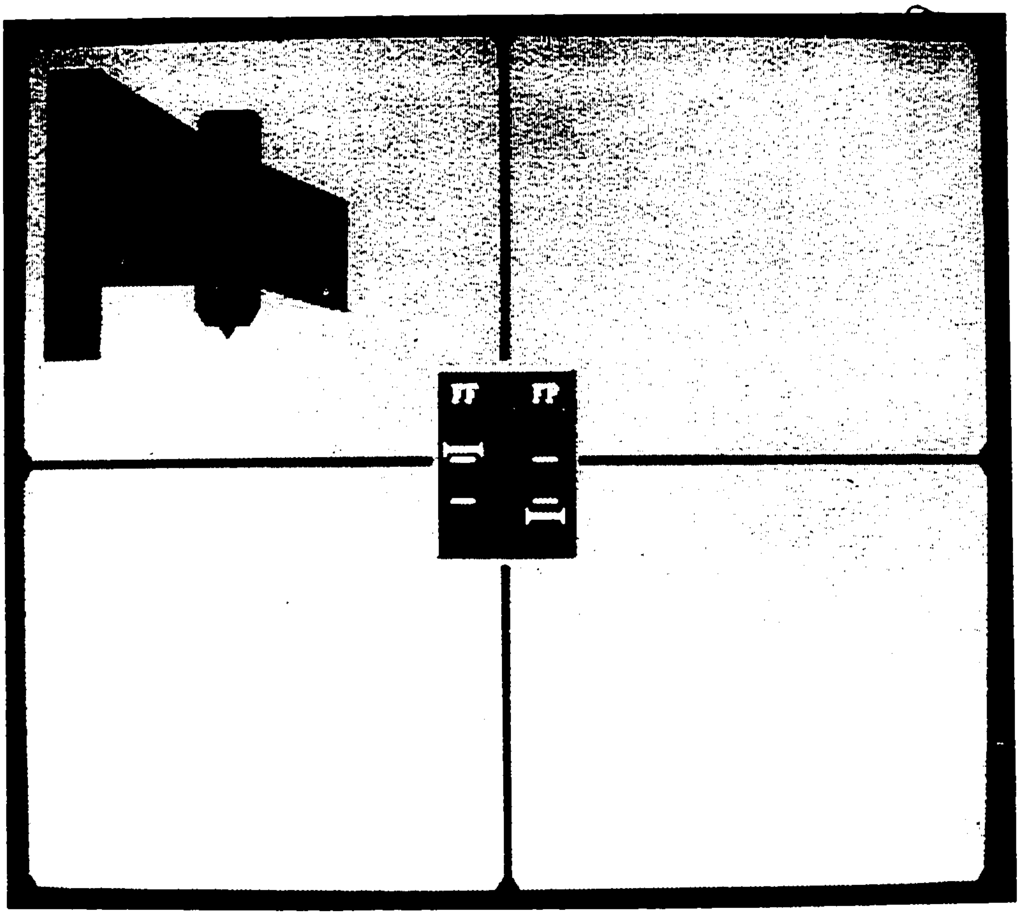
Right Fuel Leak

(Appendix Continued)



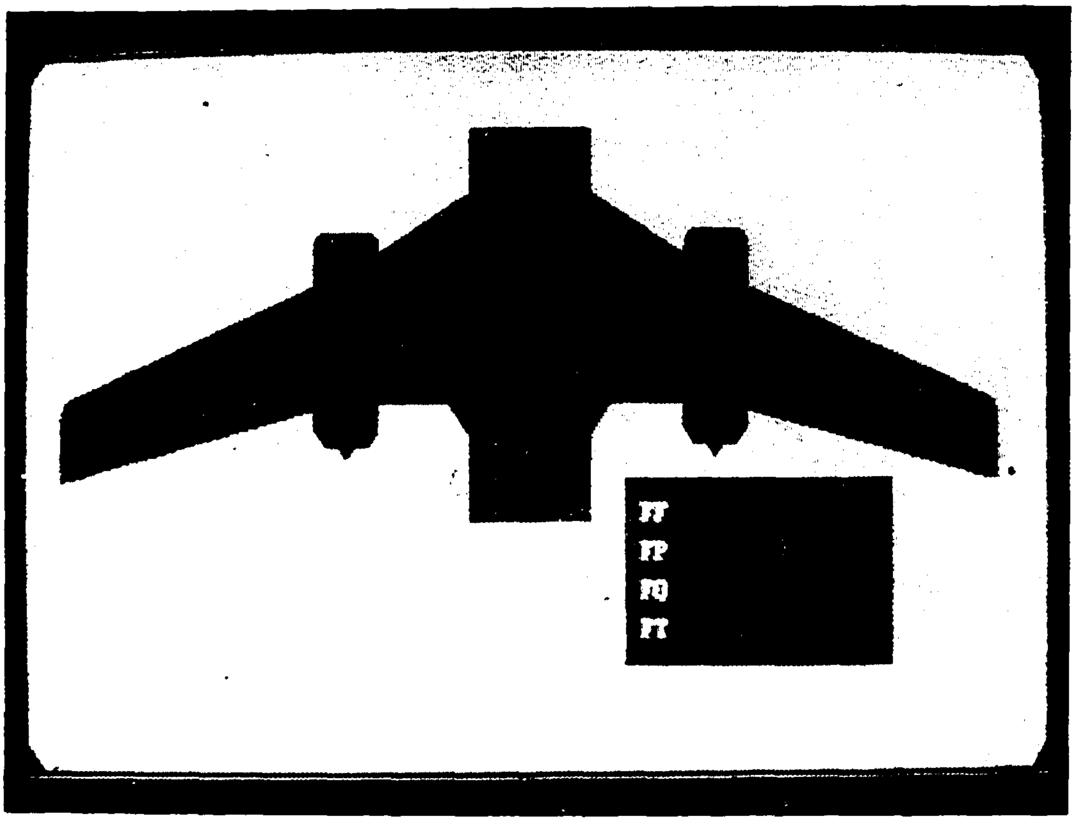
Picture - Composite Hypothesis

(Appendix Continued)



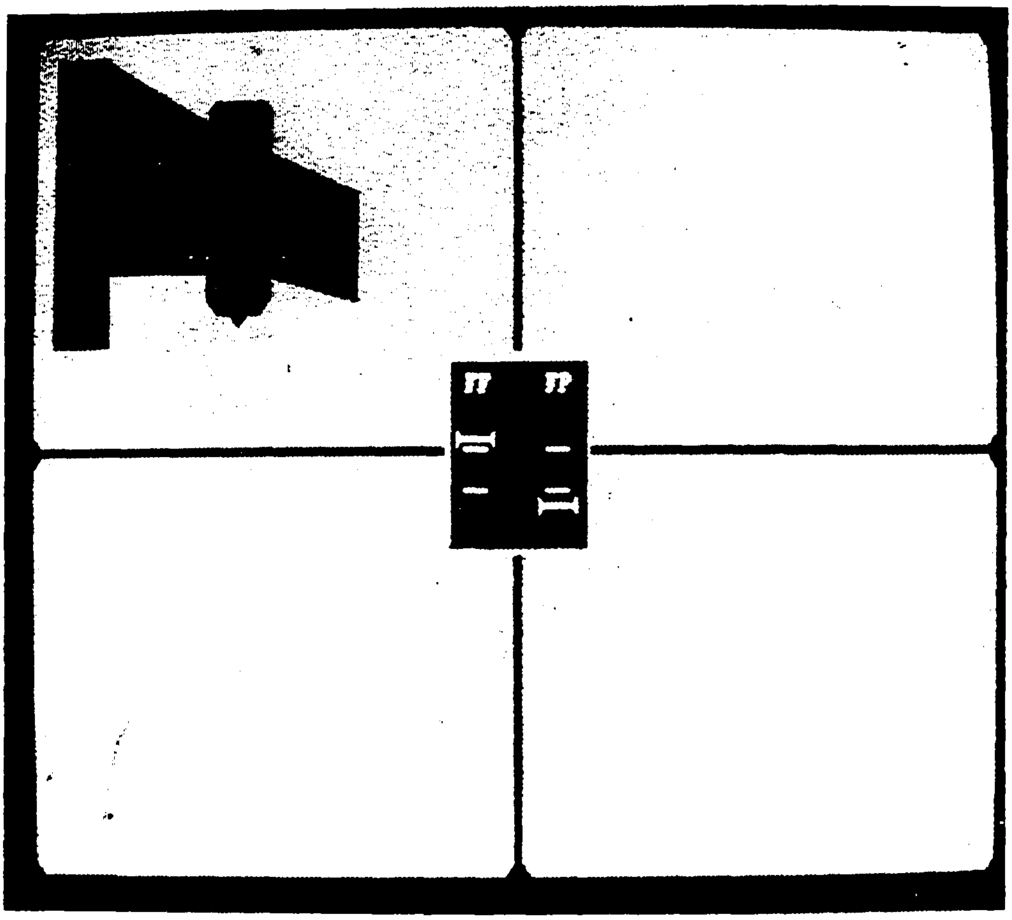
Picture - Multiple Hypotheses

(Appendix Continued)



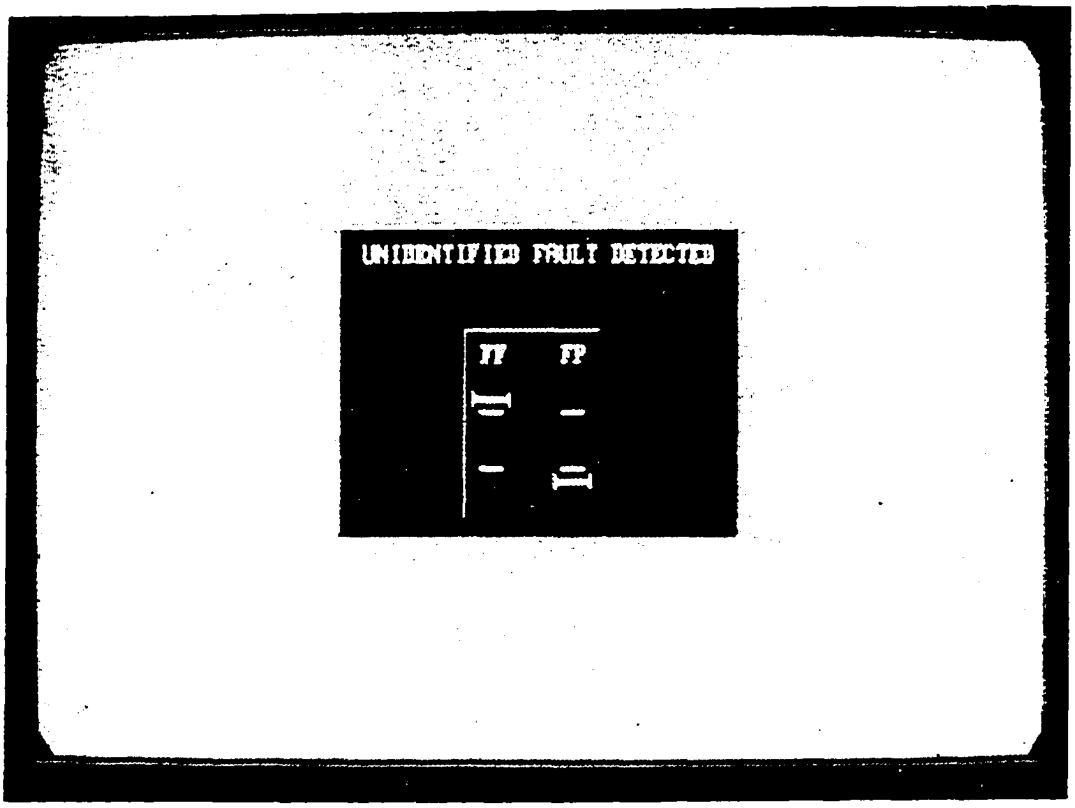
<u>Symbol - Composite Hypothesis</u>

(Appendix Continued)



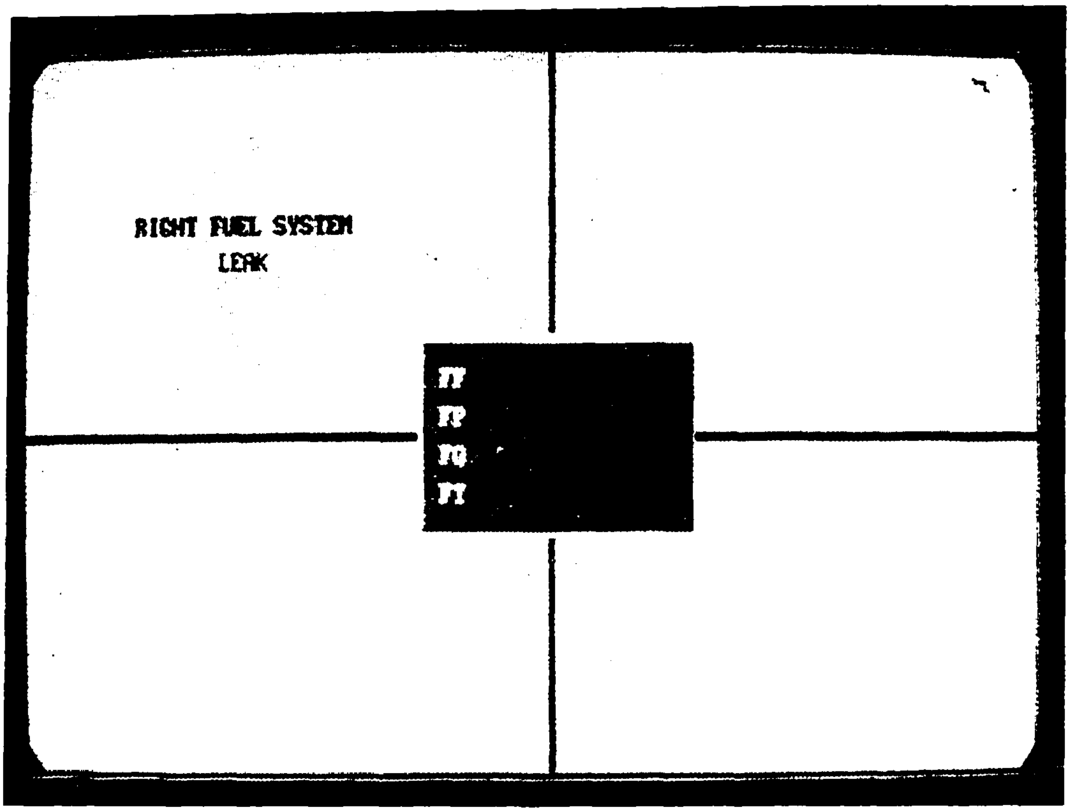
Symbol - Multiple Hypotheses

(Appendix Continued)



Text - Composite Hypothesis

(Appendix Continued)



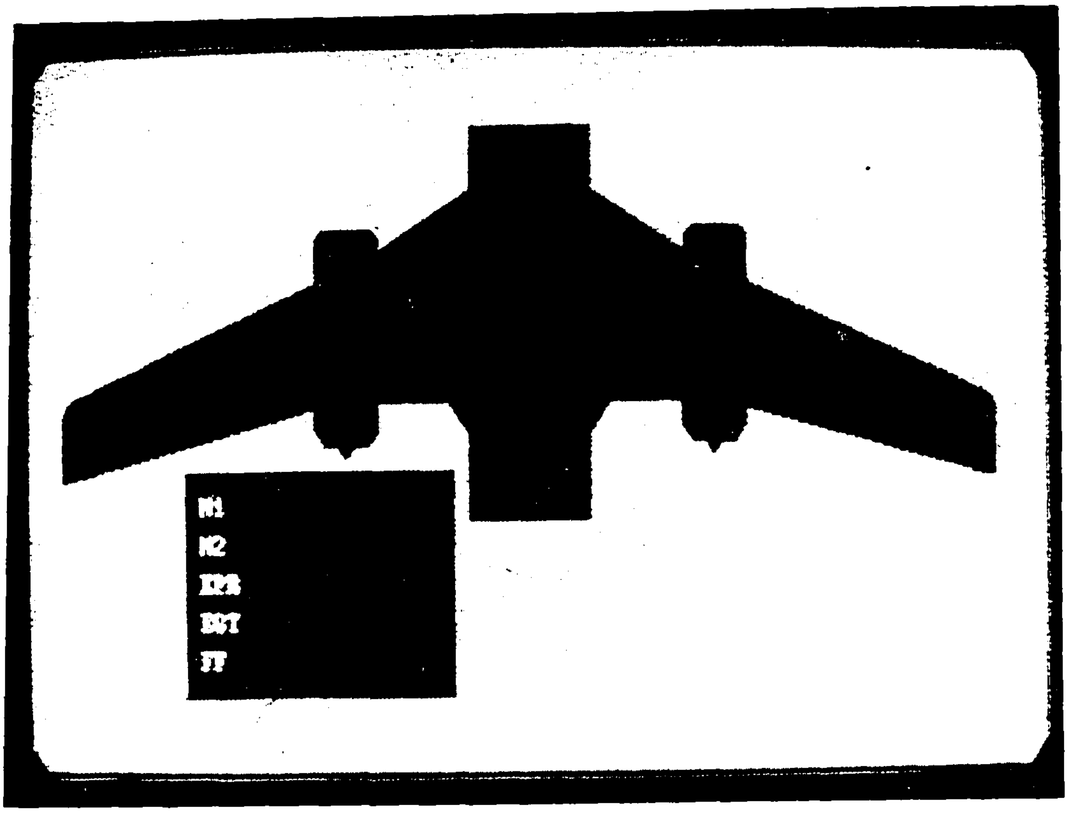
Text - Multiple Hypotheses

(Appendix Continued)

Displays for Case D

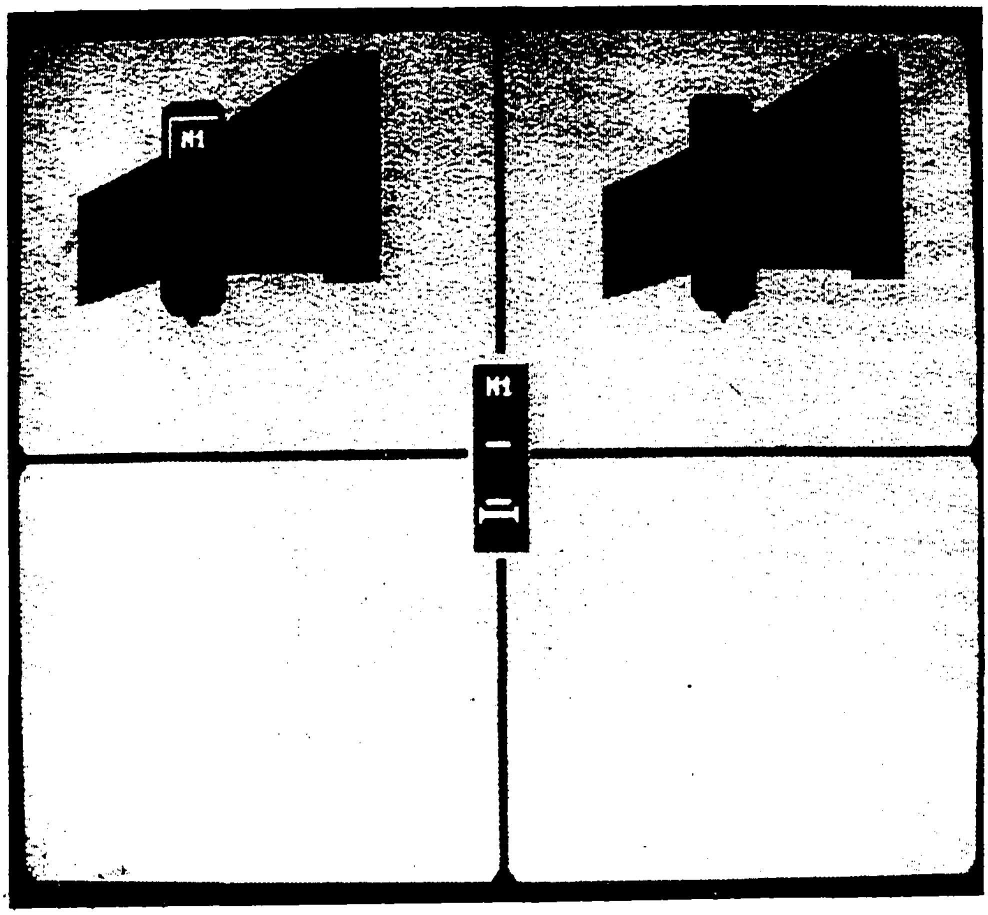
Left Engine Fault

(Appendix Continued)



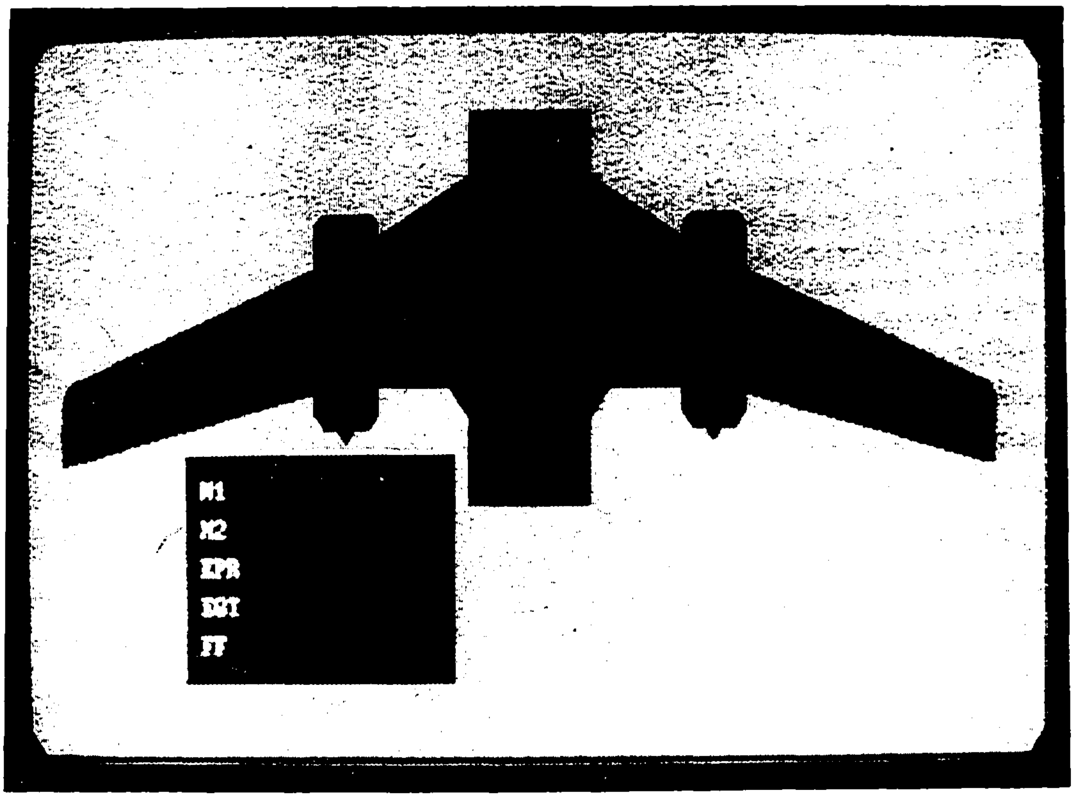
Picture - Composite Hypothesis

(Appendix Continued)



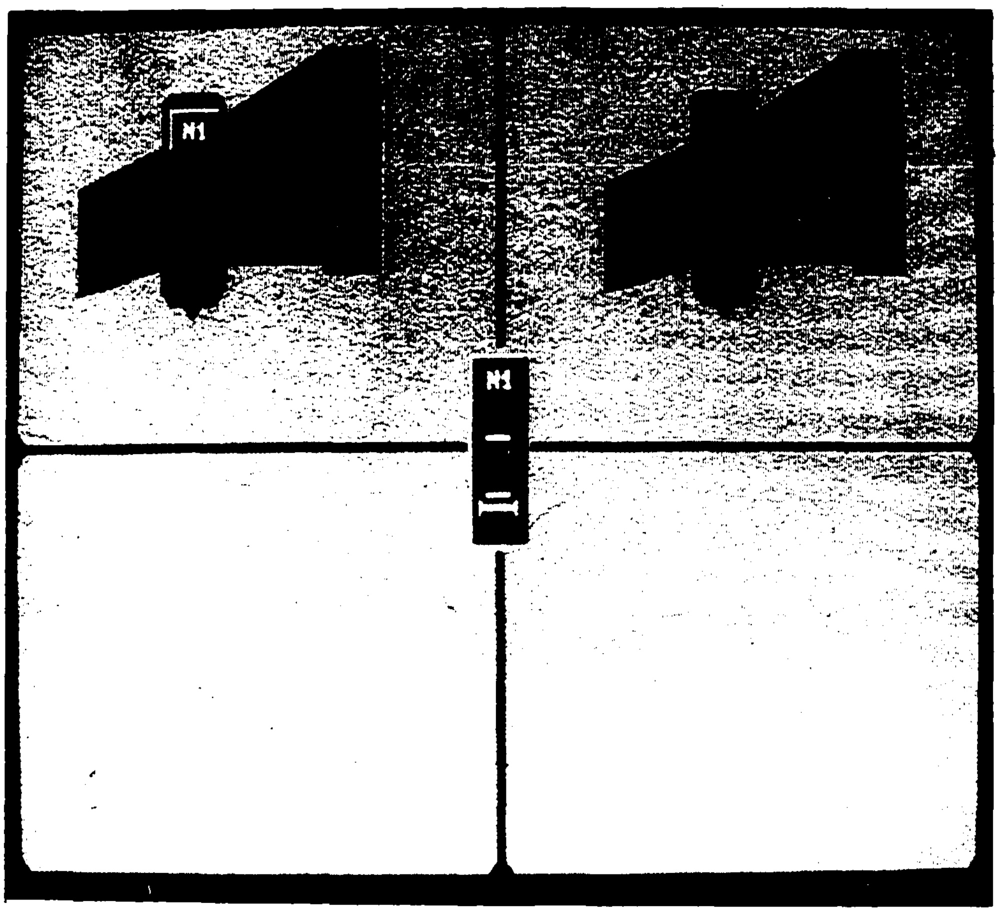
Picture - Multiple Hypotheses

(Appendix Continued)



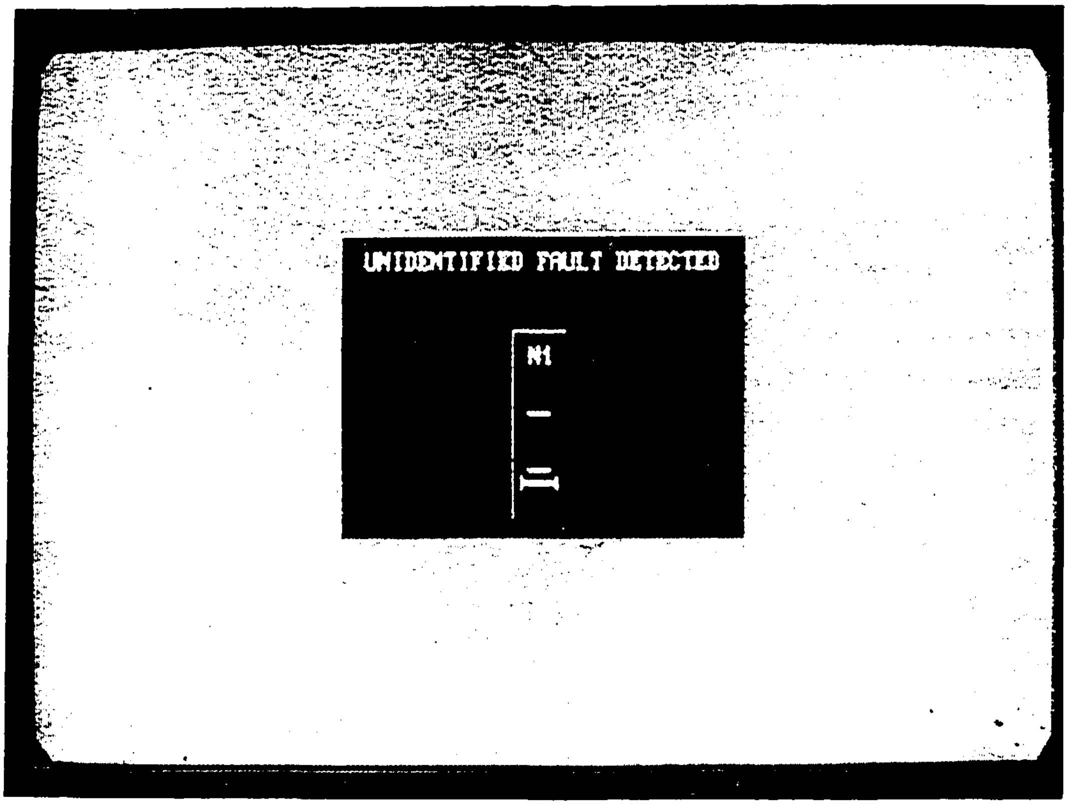
Symbol - Composite Hypothesis

(Appendix Continued)



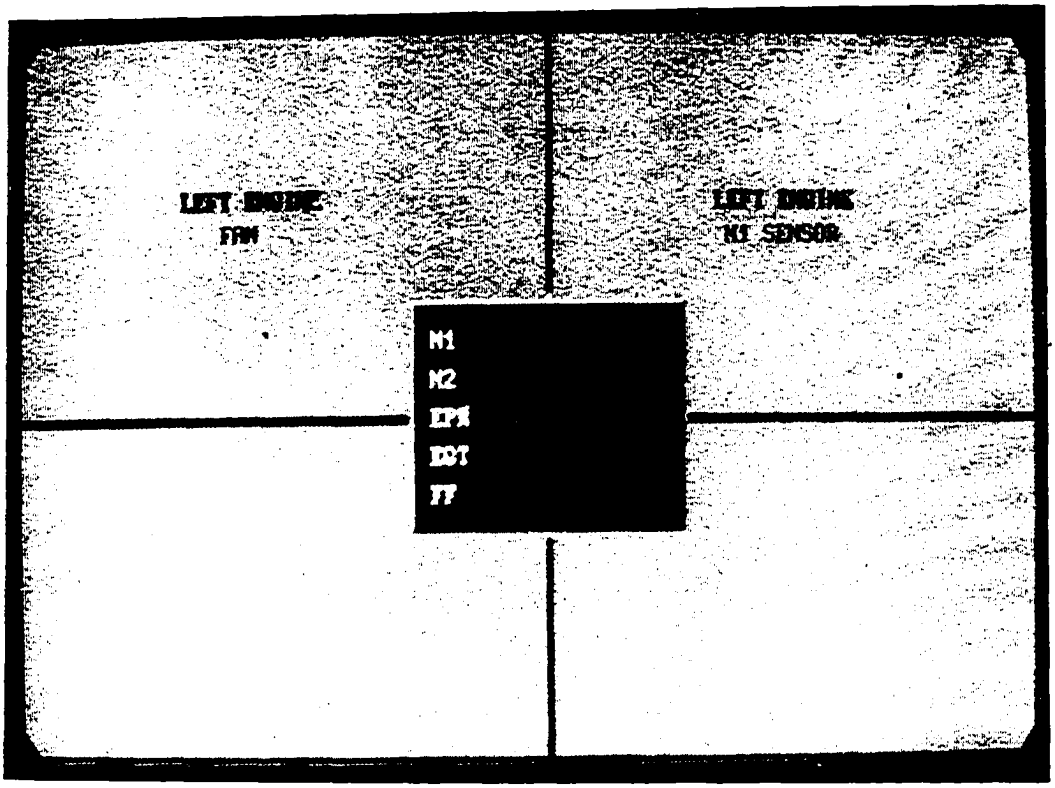
Symbol - Multiple Hypotheses

(Appendix Continued)



Text - Composite Hypothesis

(Appendix Continued)



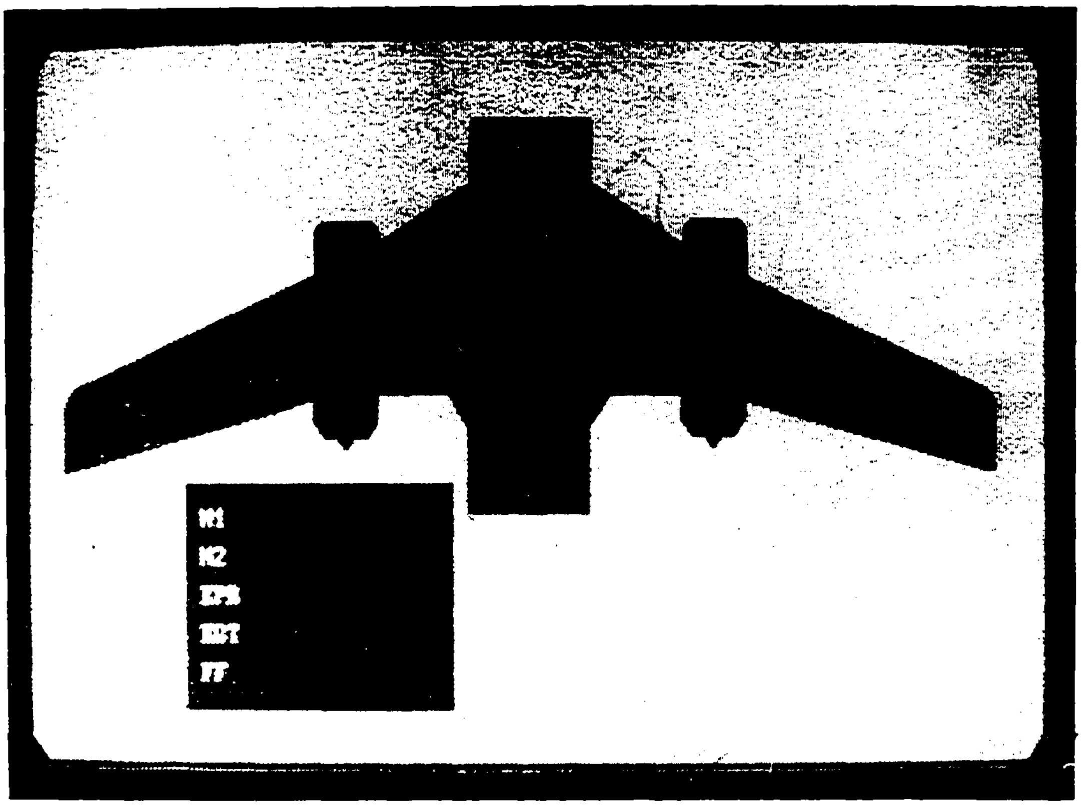
Text - Multiple Hypotheses

(Appendix Continued)

Displays for Case E

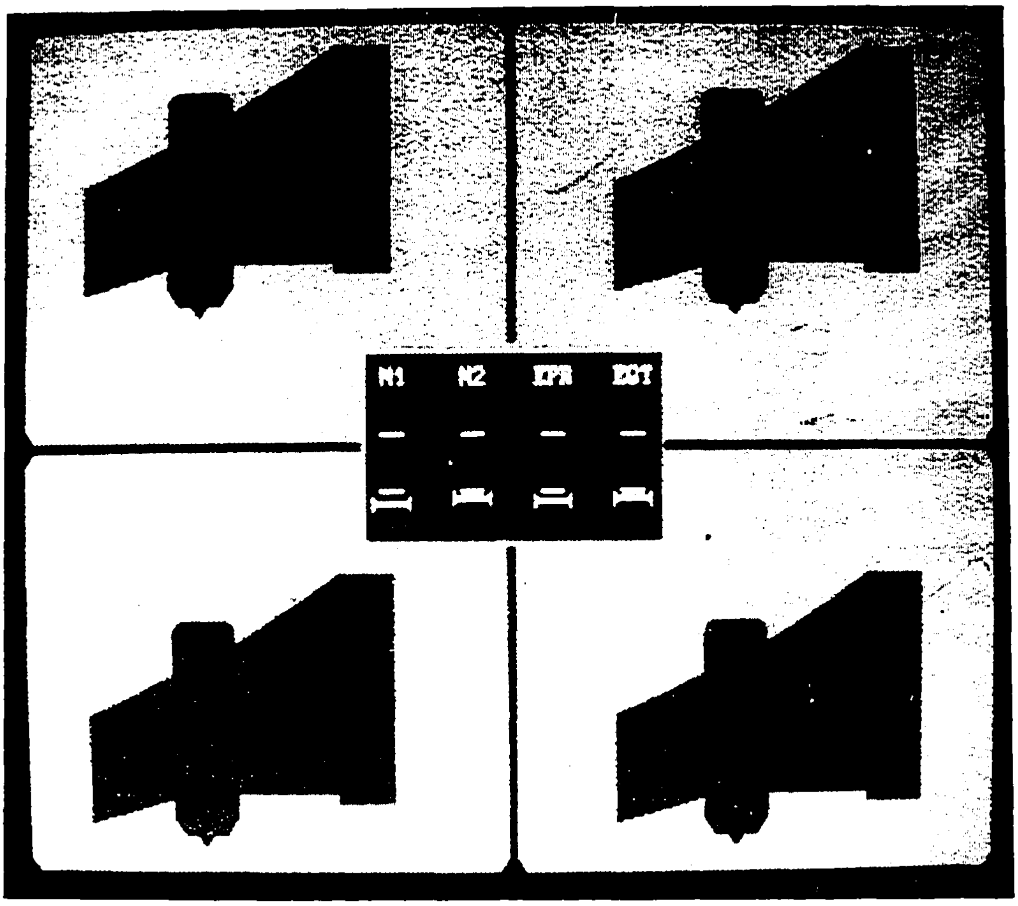
Left Engine Fault

(Appendix Continued)



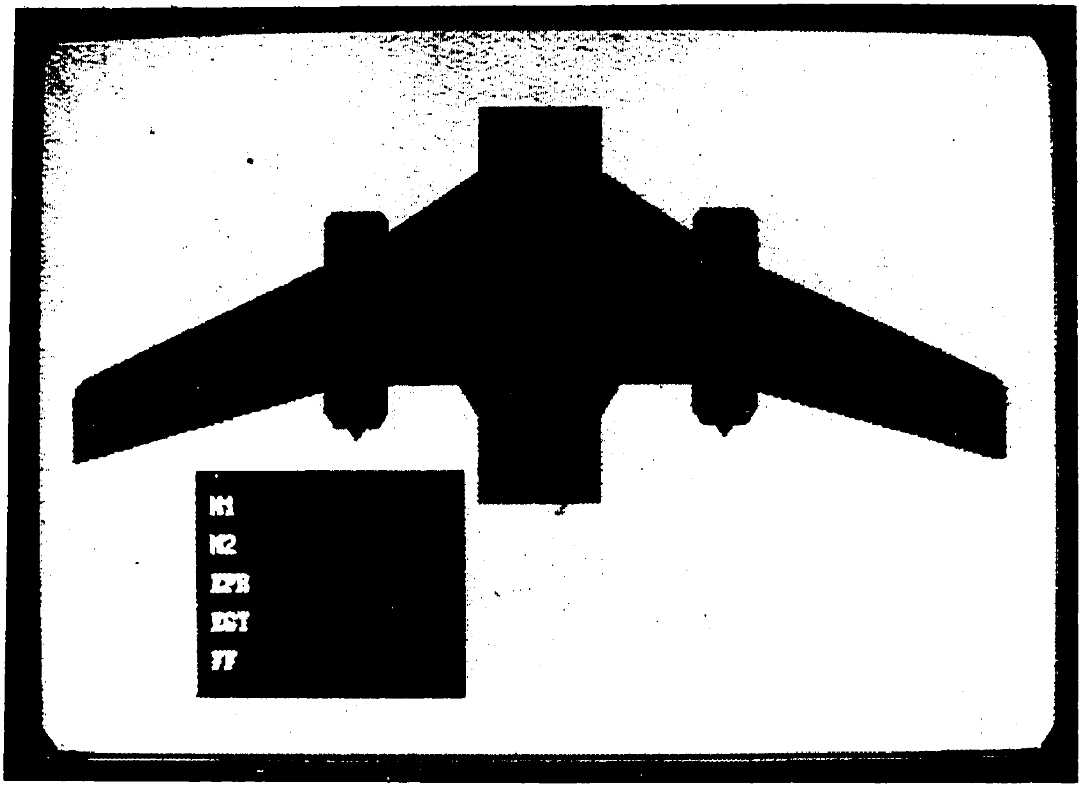
Picture - Composite Hypothesis

(Appendix Continued)



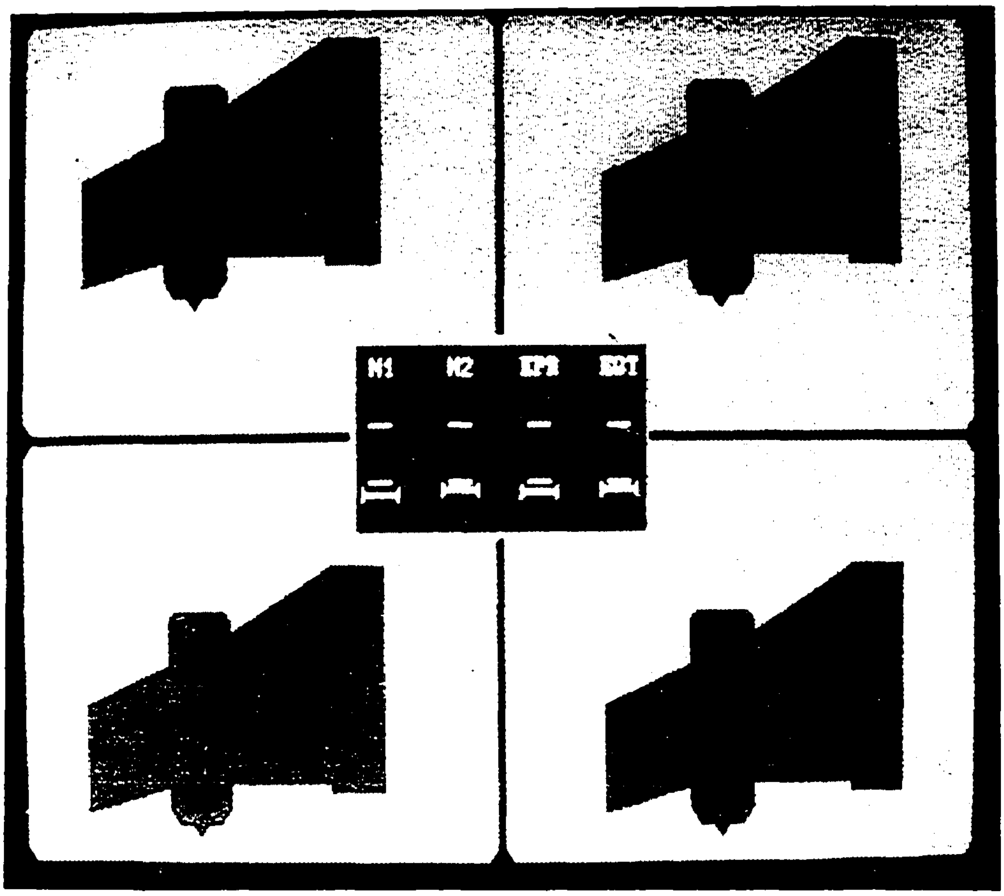
Picture - Multiple Hypotheses

(Appendix Continued)



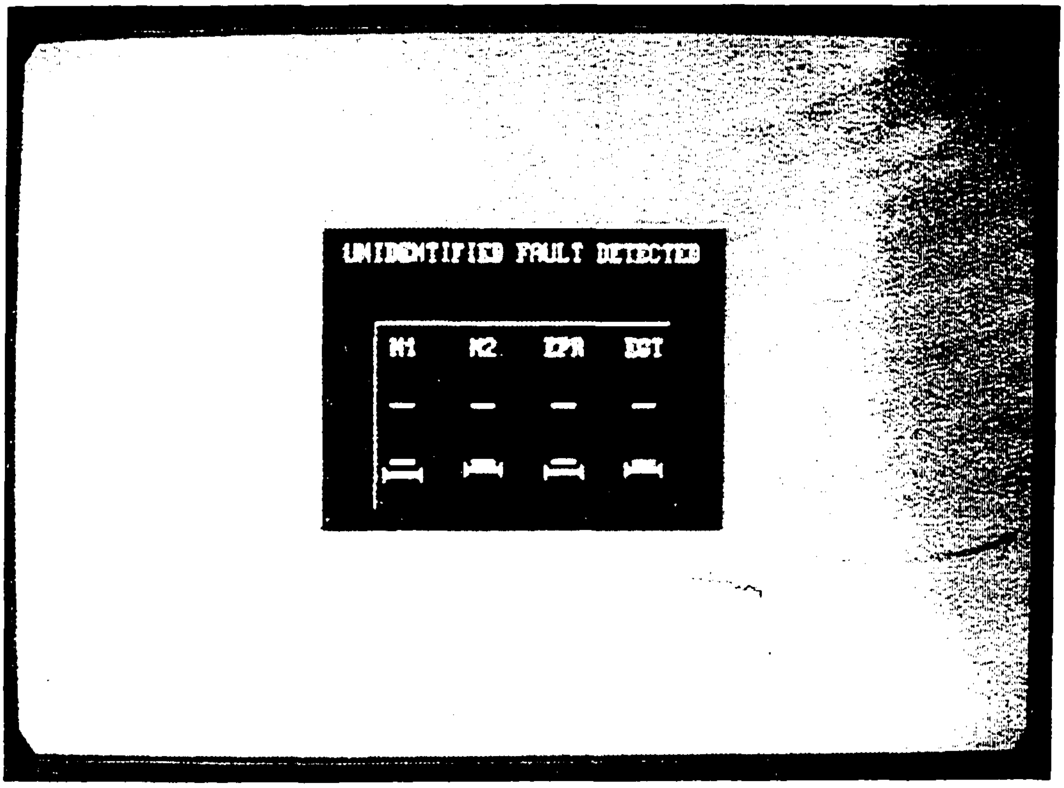
Symbol - Composite Hypothesis

(Appendix Continued)



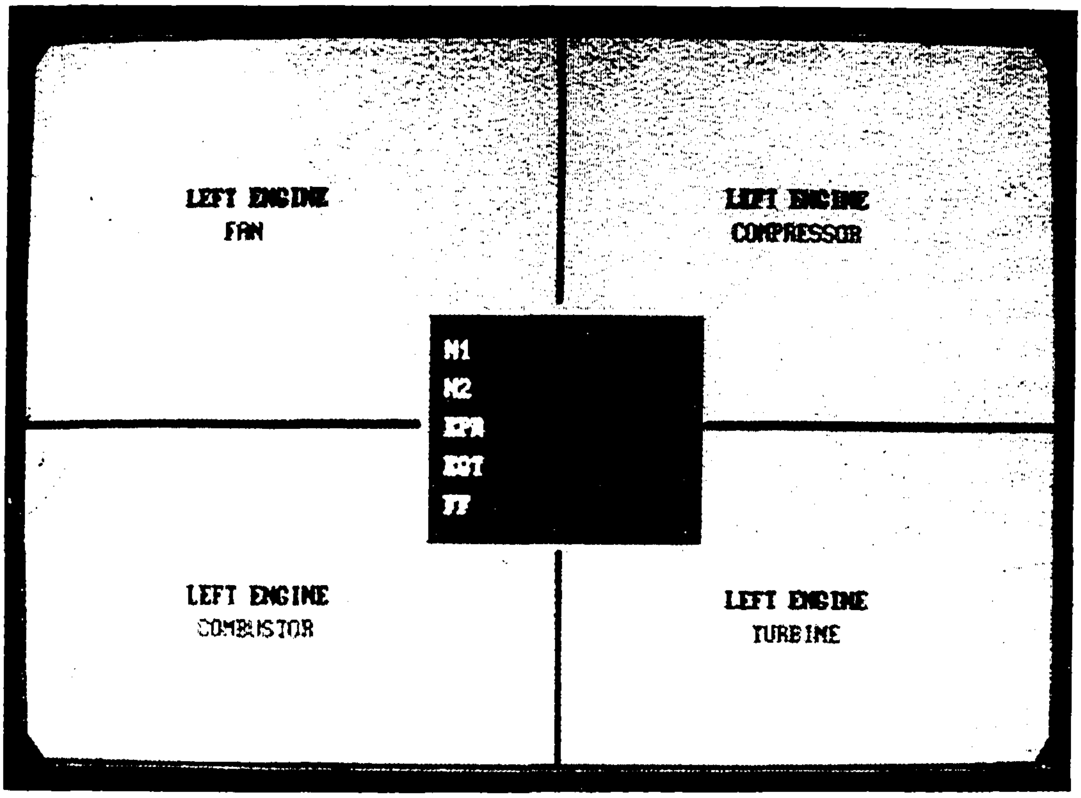
Symbol - Multiple Hypotheses

(Appendix Continued)



Text - Composite Hypothesis

(Appendix Continued)



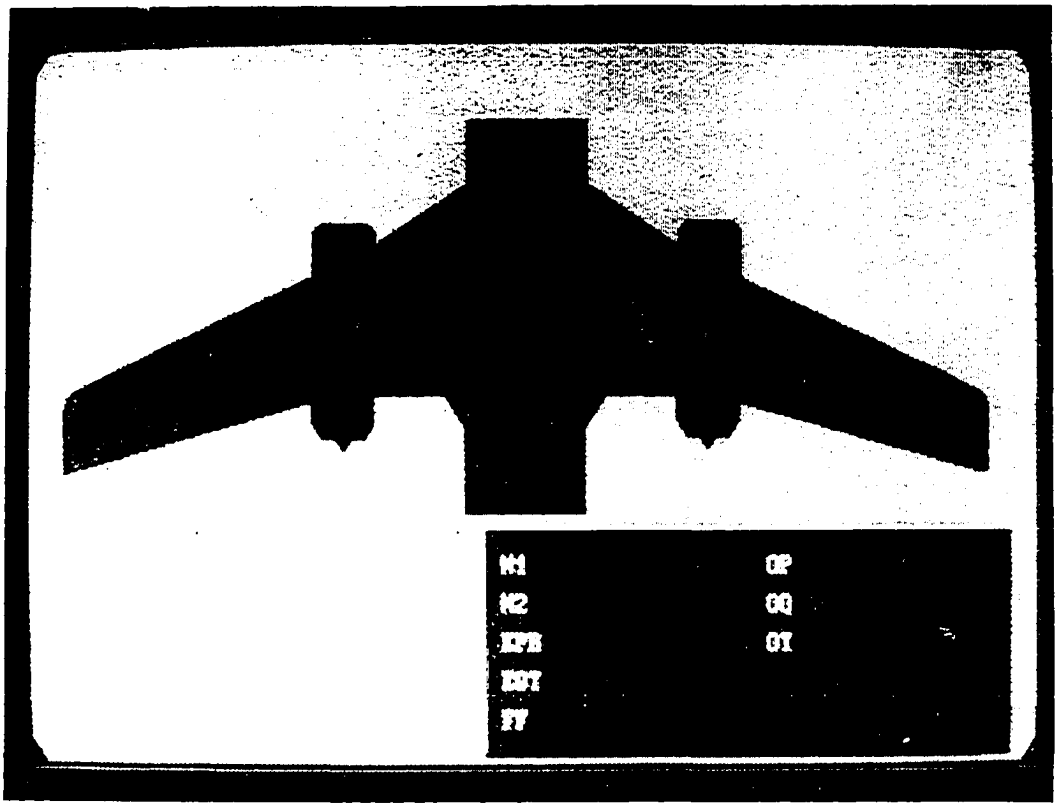
Text - Multiple Hypotheses

(Appendix Continued)

Displays for Case F

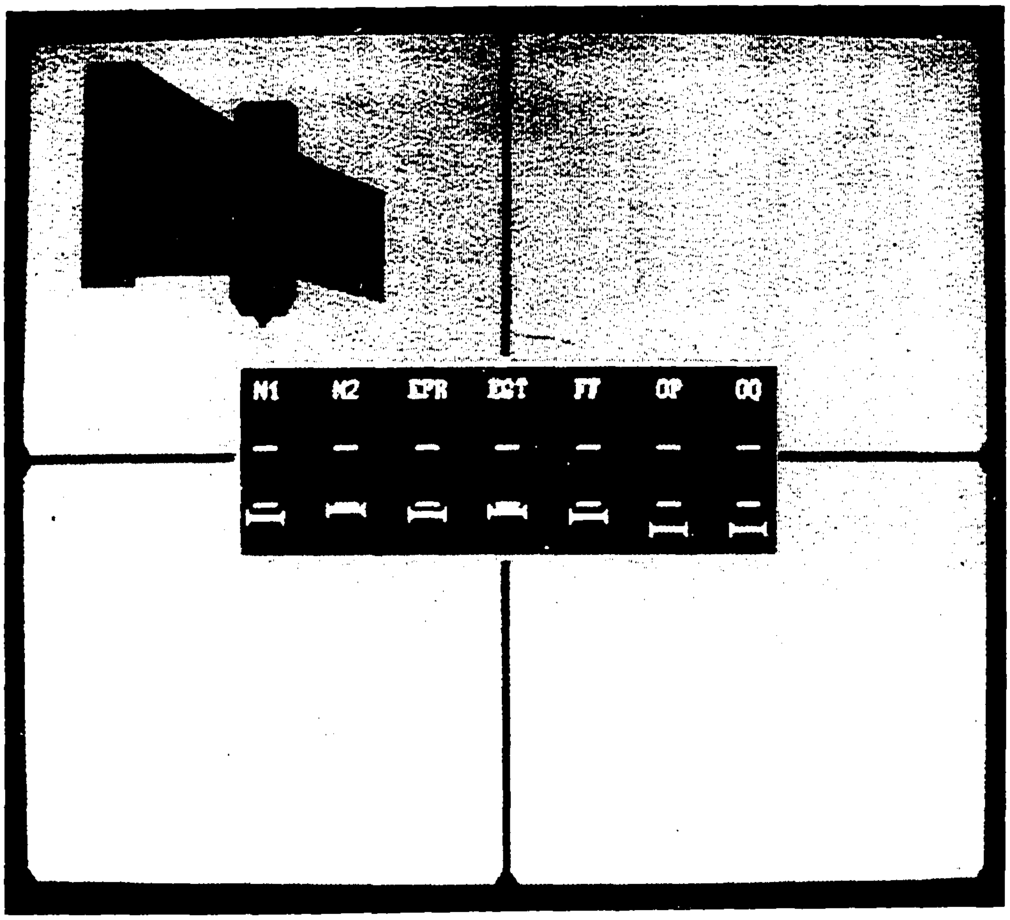
Left Engine Fault and Oil Leak

(Appendix Continued)



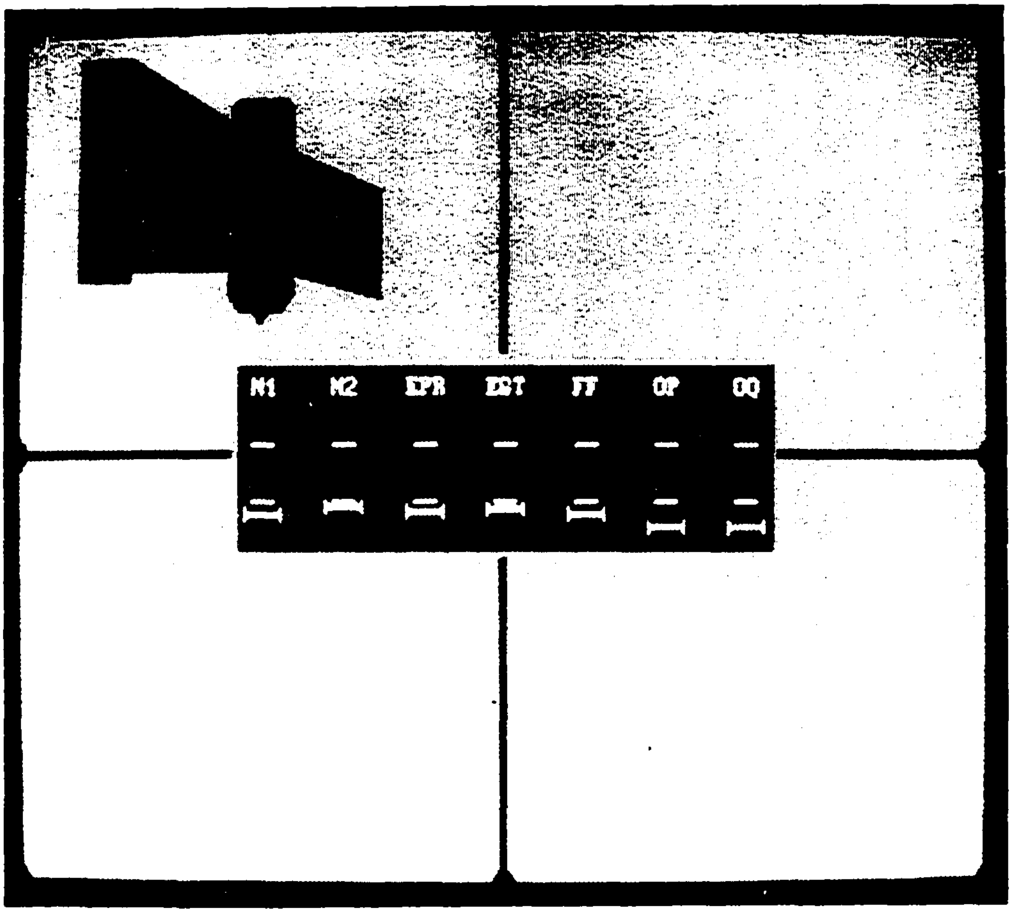
Picture - Composite Hypothesis

(Appendix Continued)



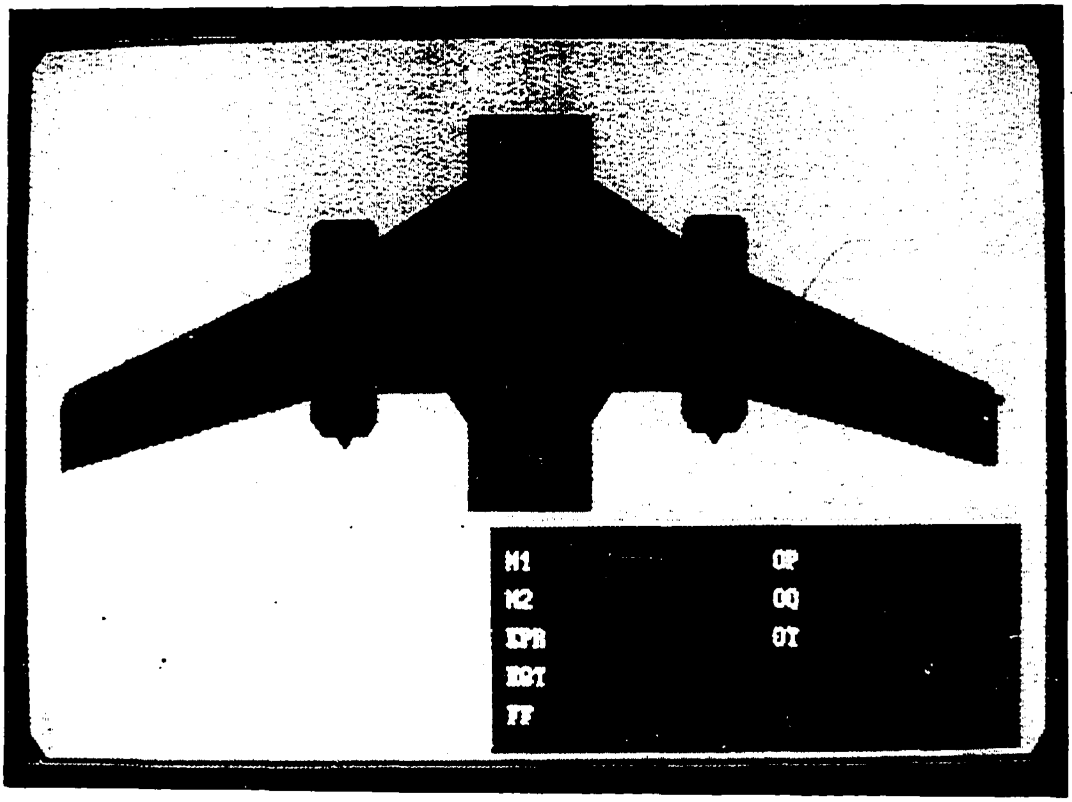
Picture - Multiple Hypotheses

(Appendix Continued)



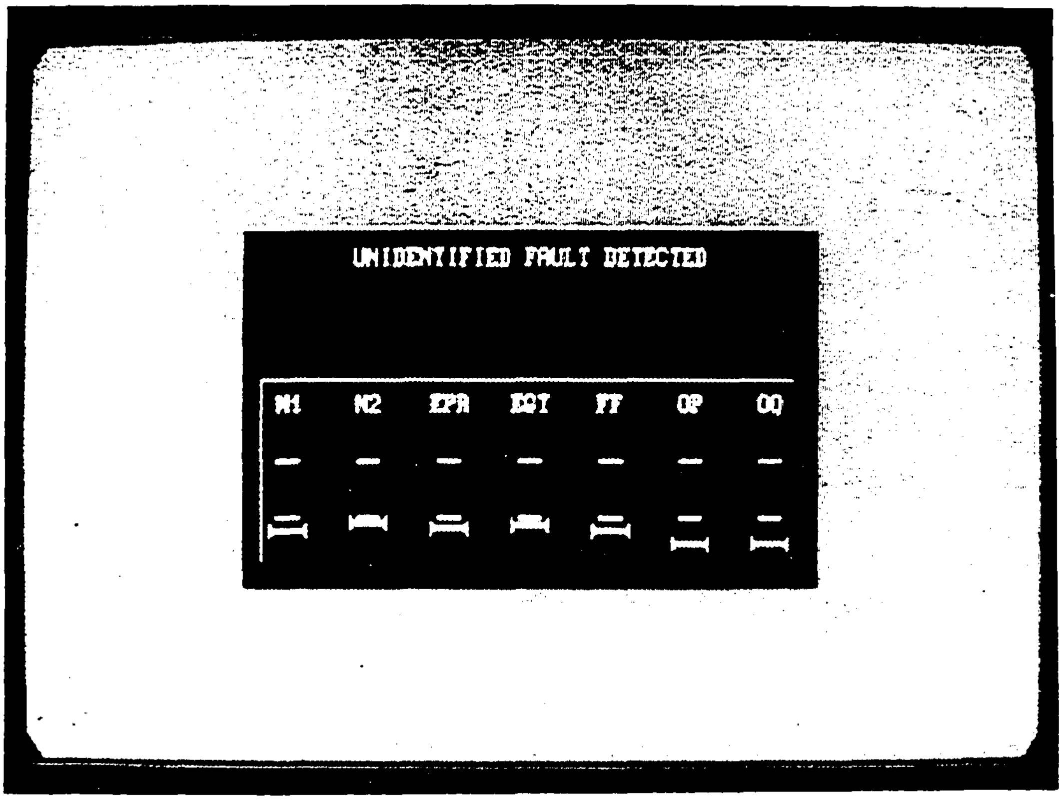
Symbol - Composite Hypothesis

(Appendix Continued)



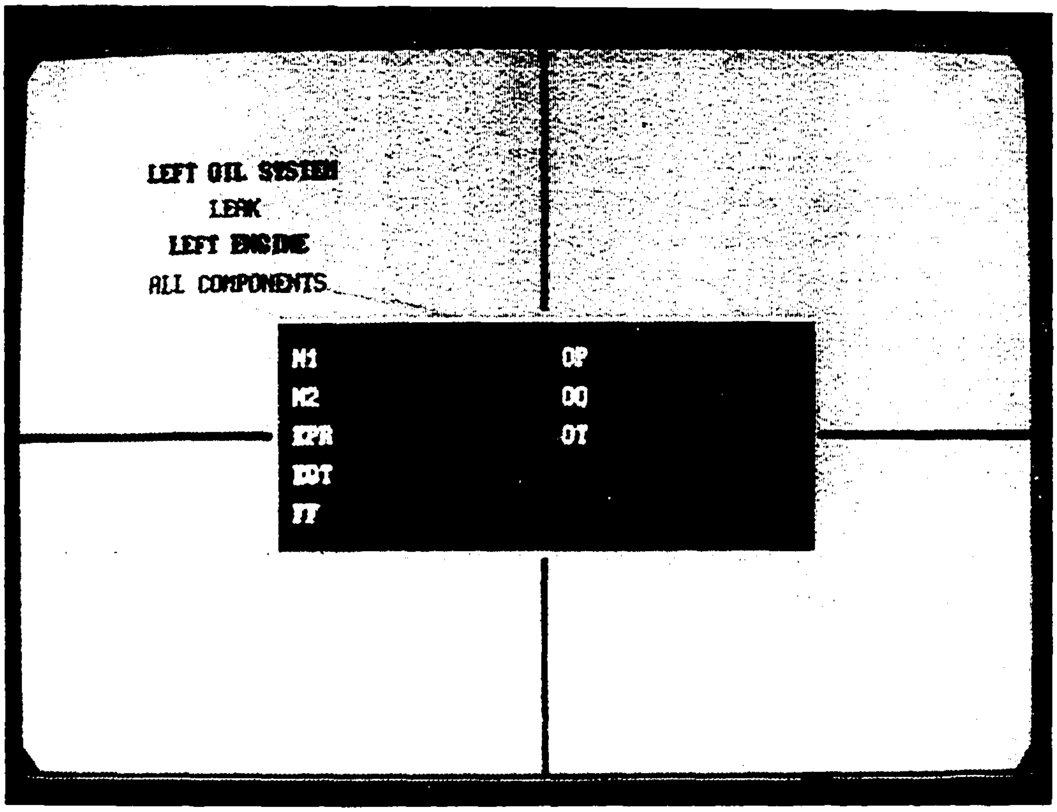
Symbol - Multiple Hypotheses

(Appendix Continued)



Text - Composite Hypothesis

(Appendix Continued)



<u>Text - Multiple Hypotheses</u>

(Appendix Continued)

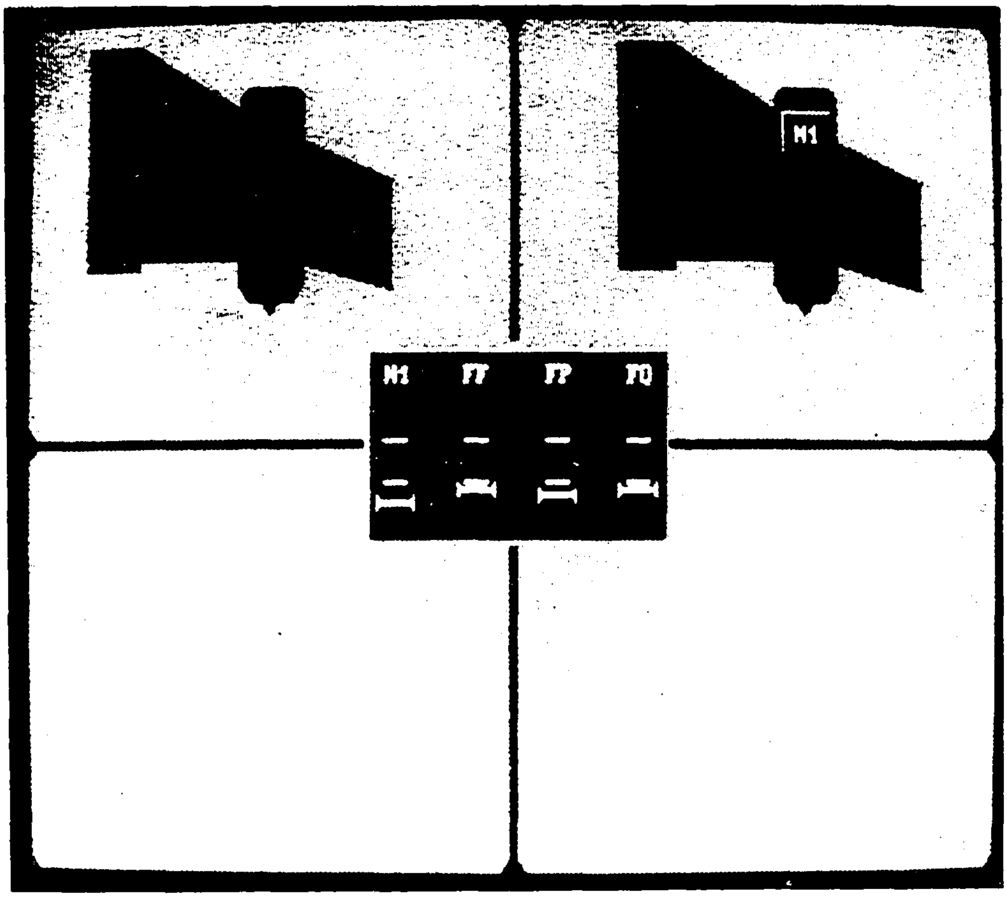
Displays for Case G

Right Engine Fault and Left Fuel Leak

(Appendix Continued)

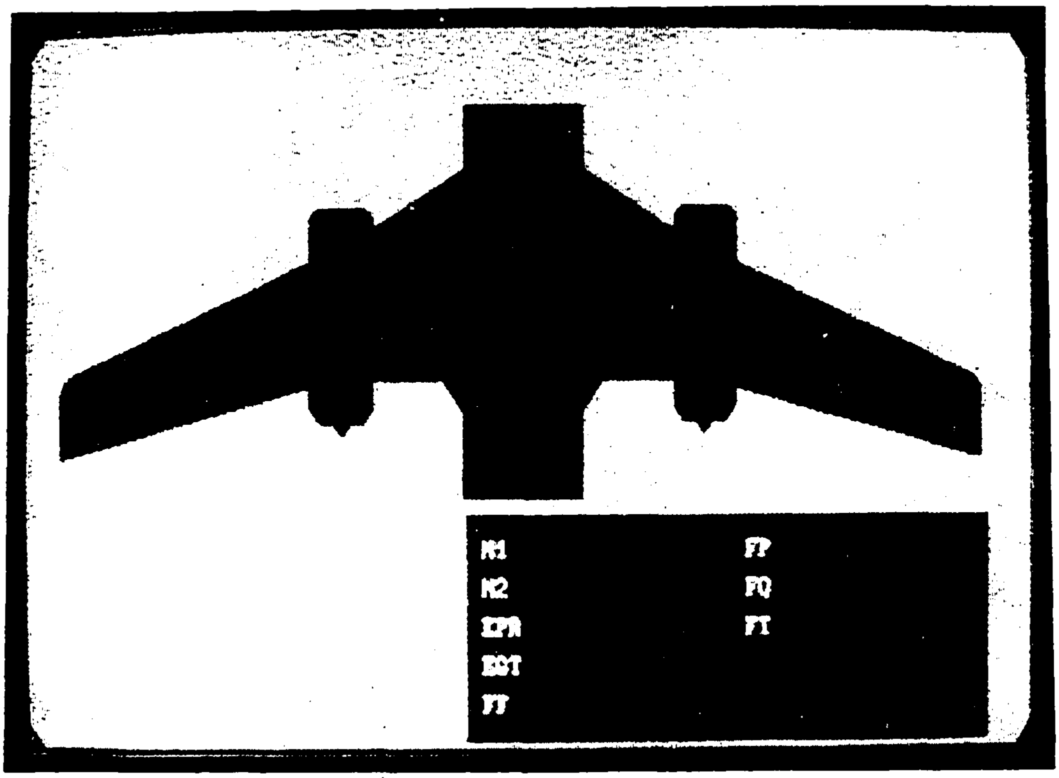
Picture - Composite Hypothesis

(Appendix Continued)



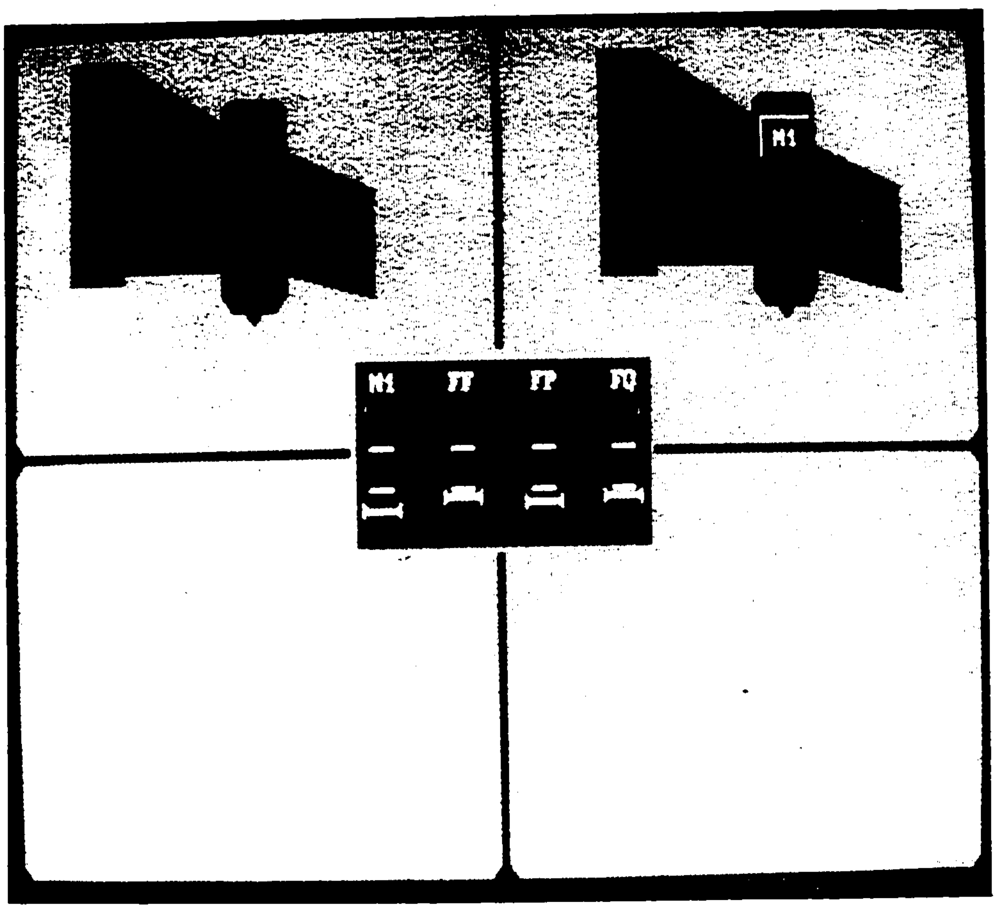
Picture - Multiple Hypotheses

(Appendix Continued)



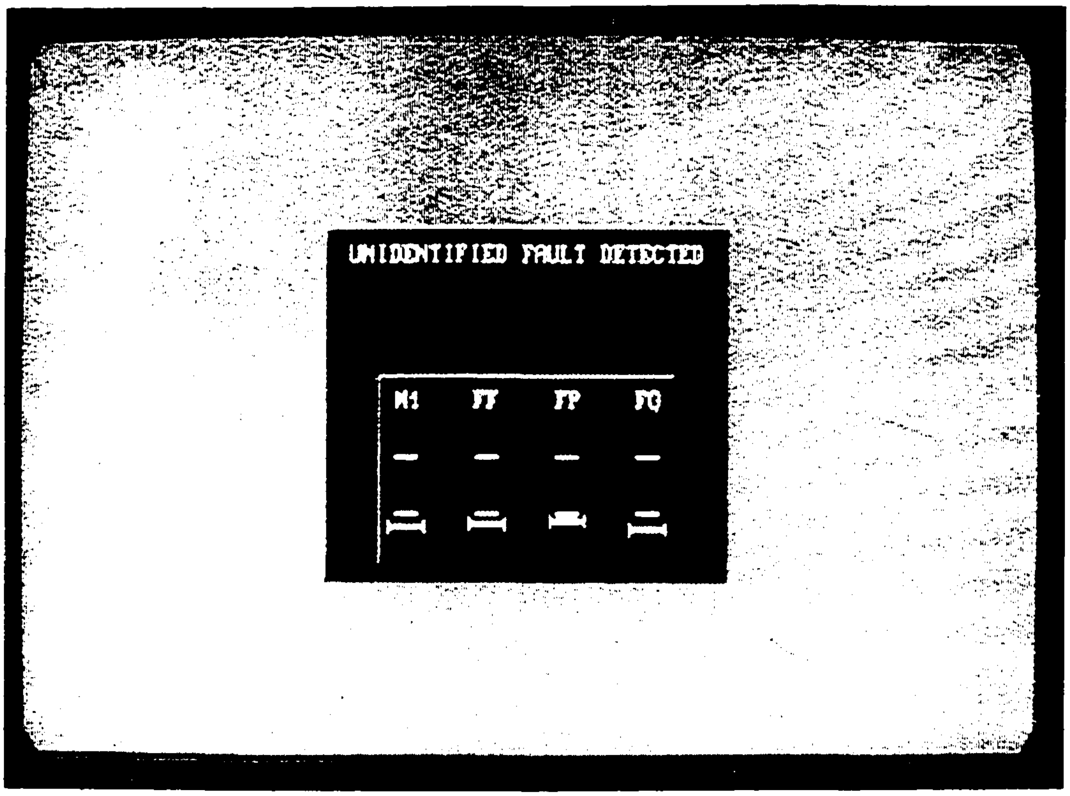
Symbol - Composite Hypothesis

(Appendix Continued)



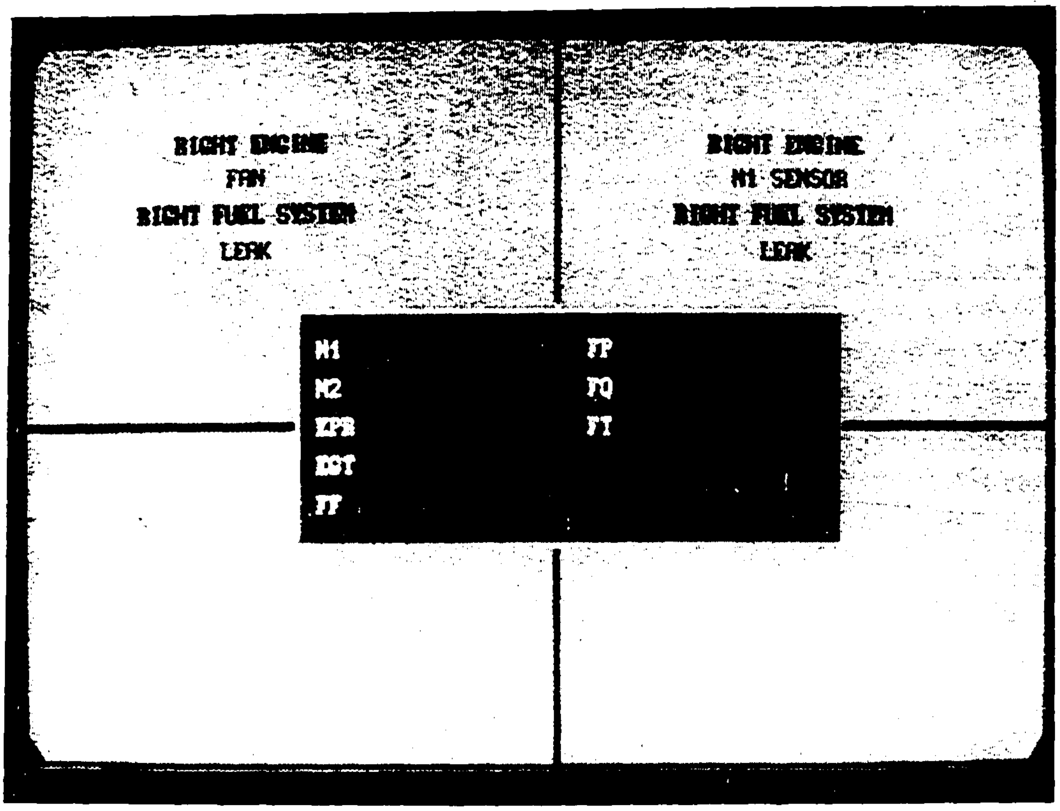
Symbol - Multiple Hypotheses

(Appendix Continued)



Text - Composite Hypothesis

(Appendix Continued)



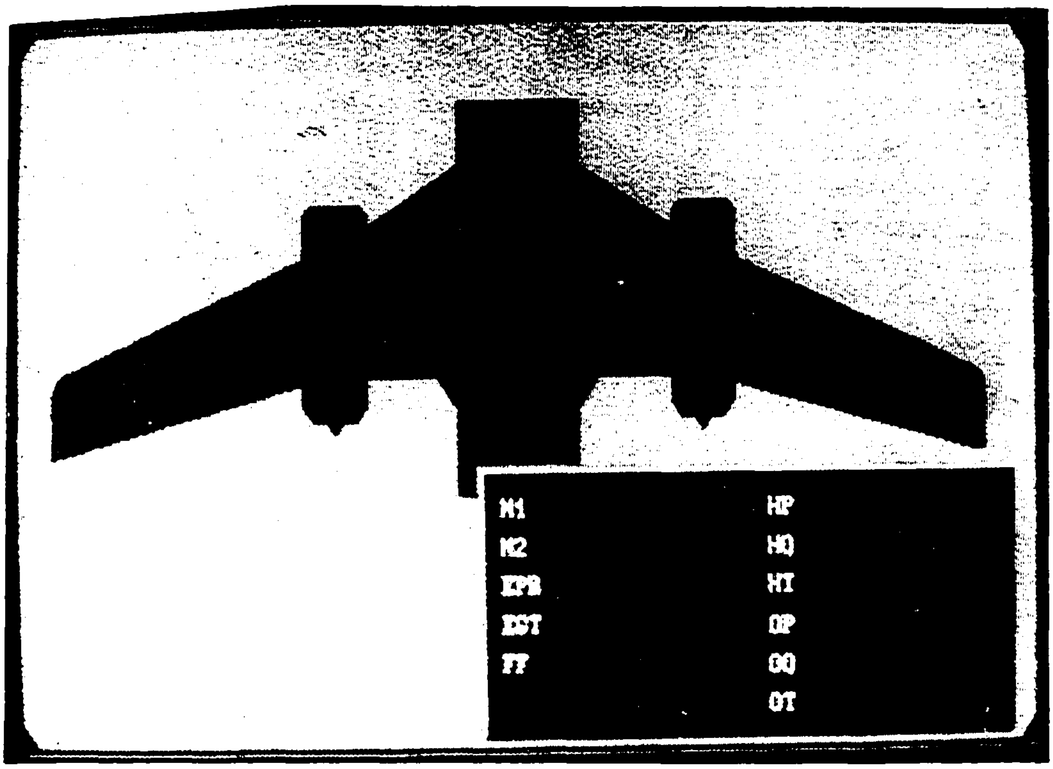
Text - Multiple Hypotheses

(Appendix Continued)

Displays for Case H

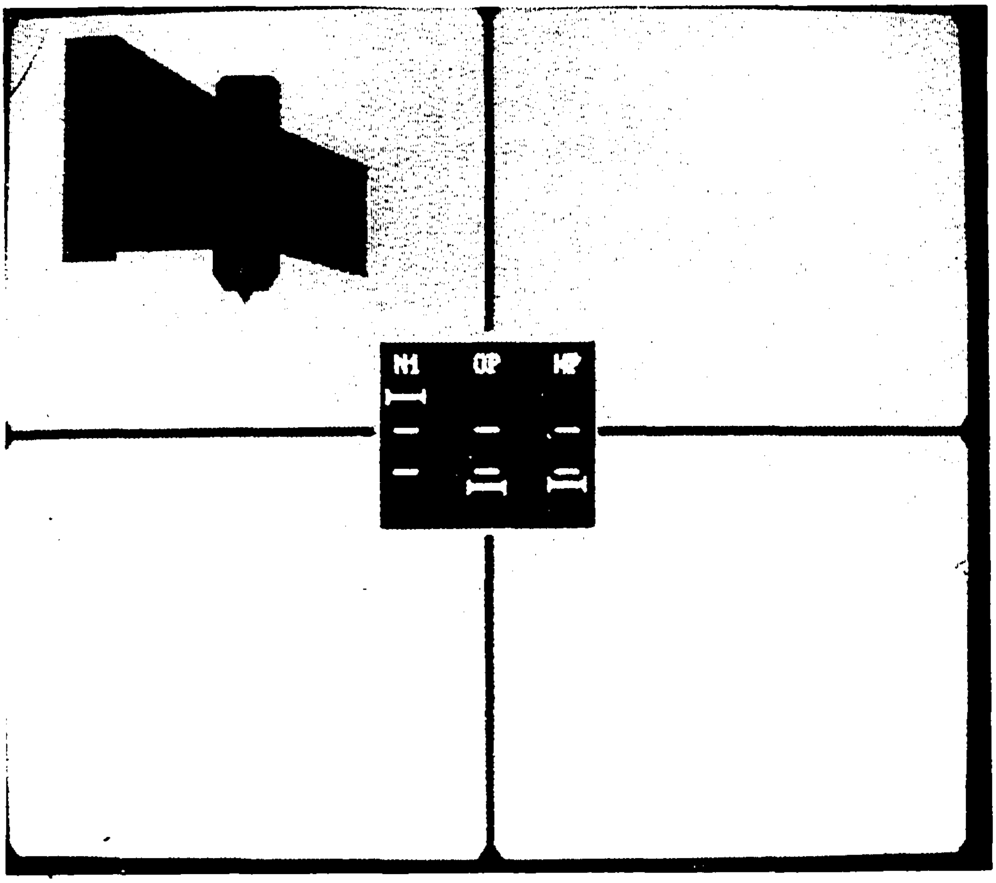
Right Engine Fault, Hydraulic Leak, and Oil Leak

(Appendix Continued)



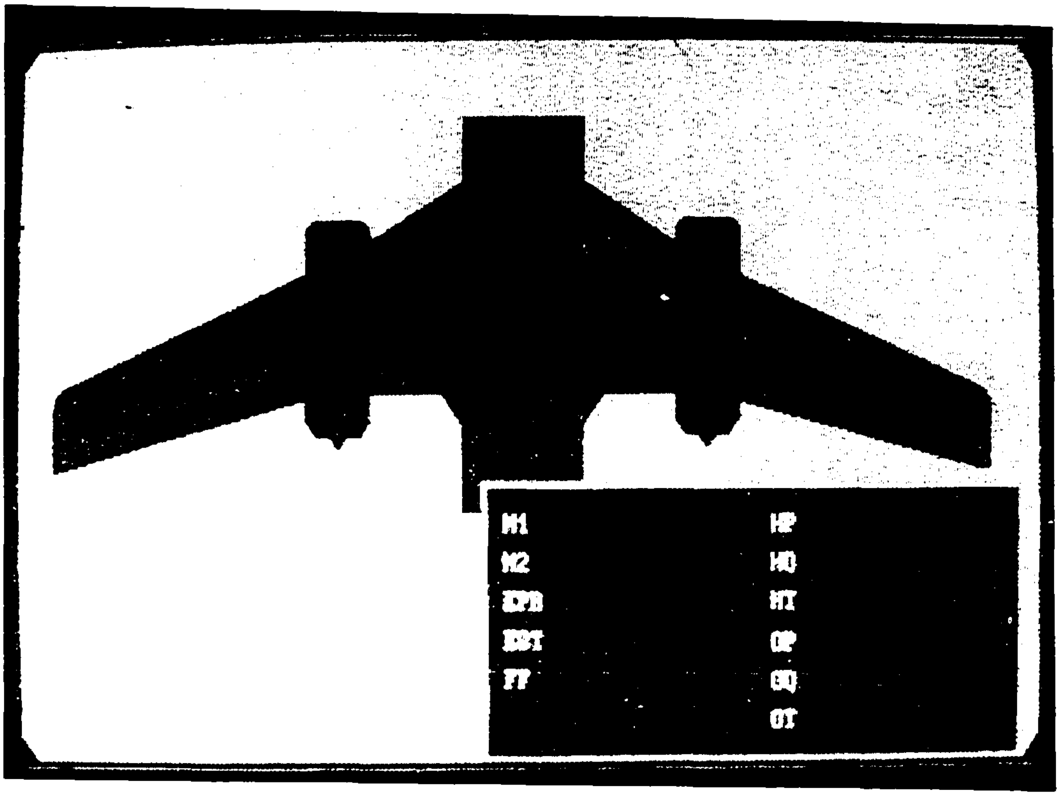
Picture - Composite Hypothesis

(Appendix Continued)



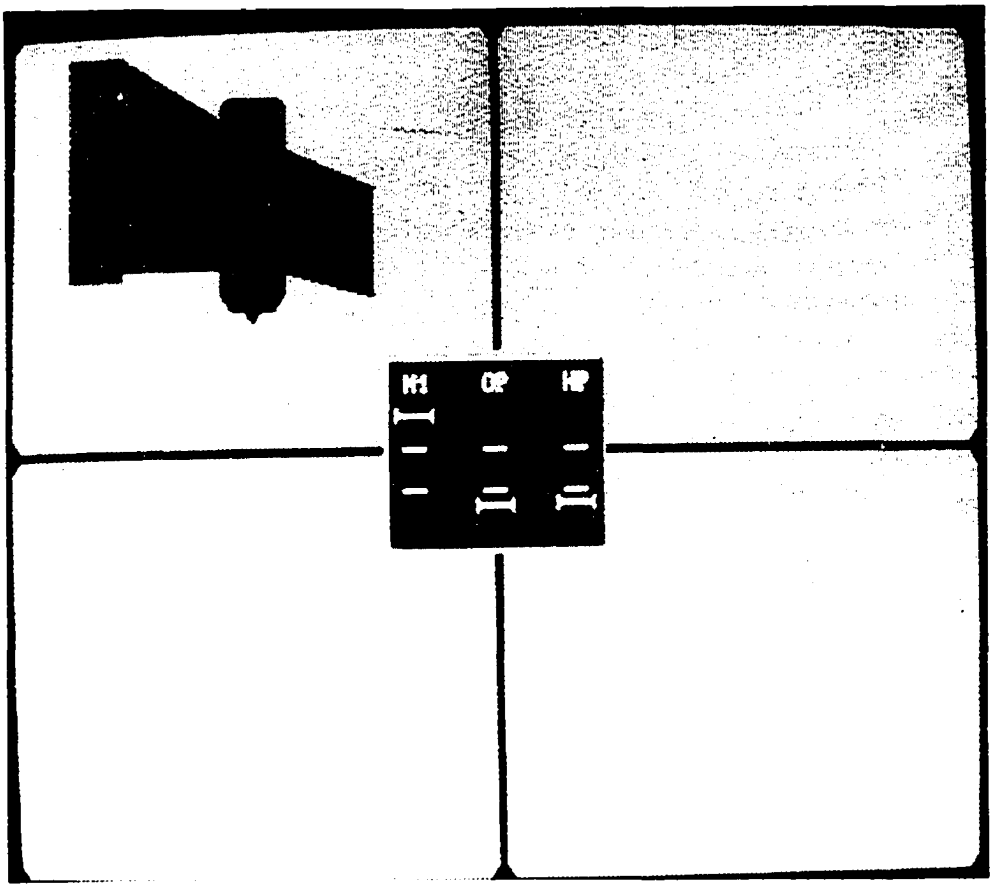
Picture - Multiple Hypotheses

(Appendix Continued)



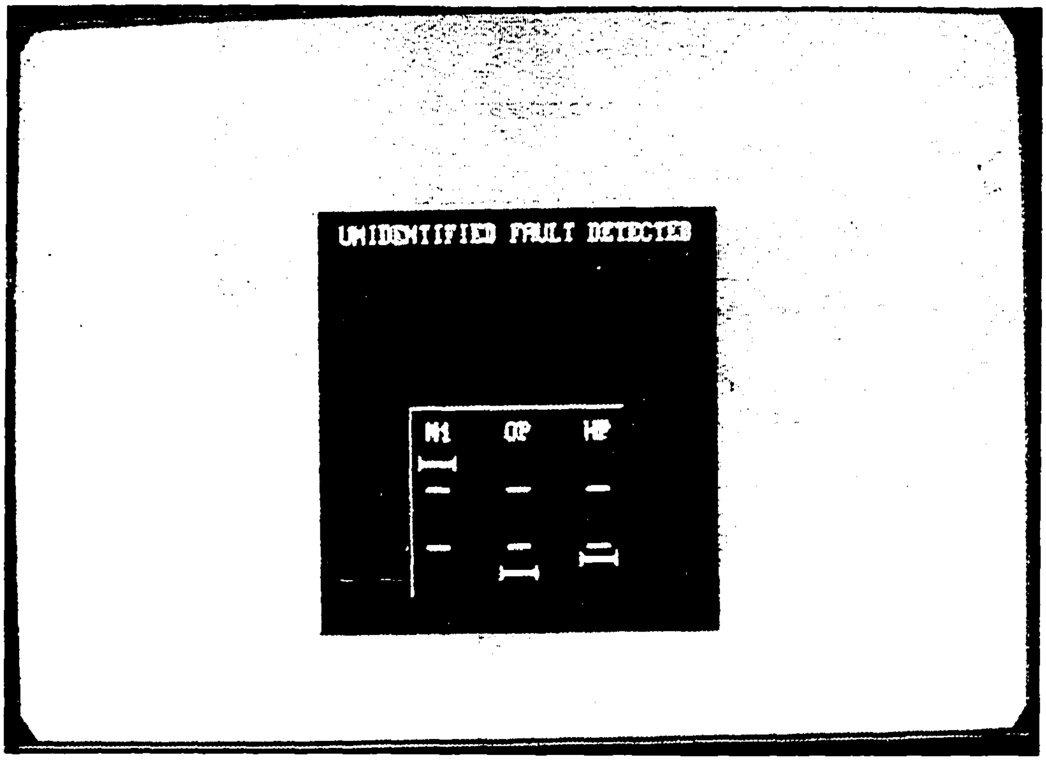
Symbol - Composite Hypothesis

(Appendix Continued)



Symbol - Multiple Hypotheses

(Appendix Continued)



Text - Composite Hypothesis

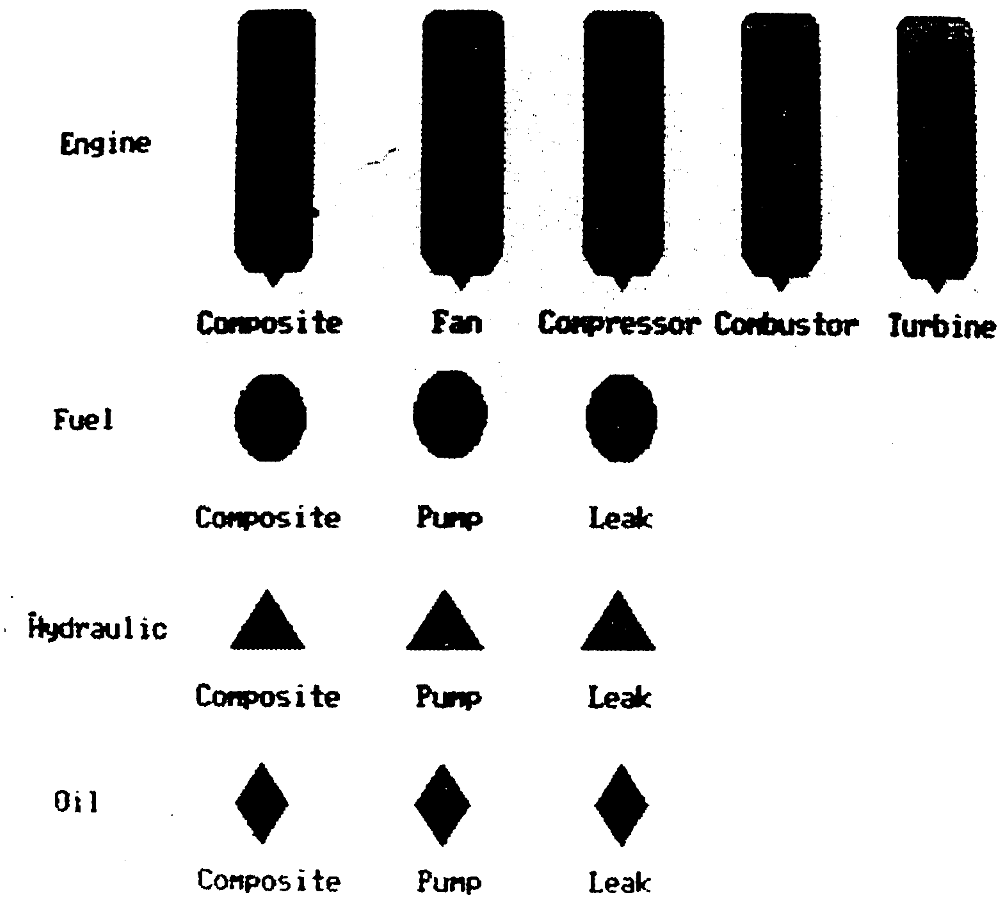
(Appendix Continued)

Text - Multiple Hypotheses

Appendix B

Symbols Used in Displays

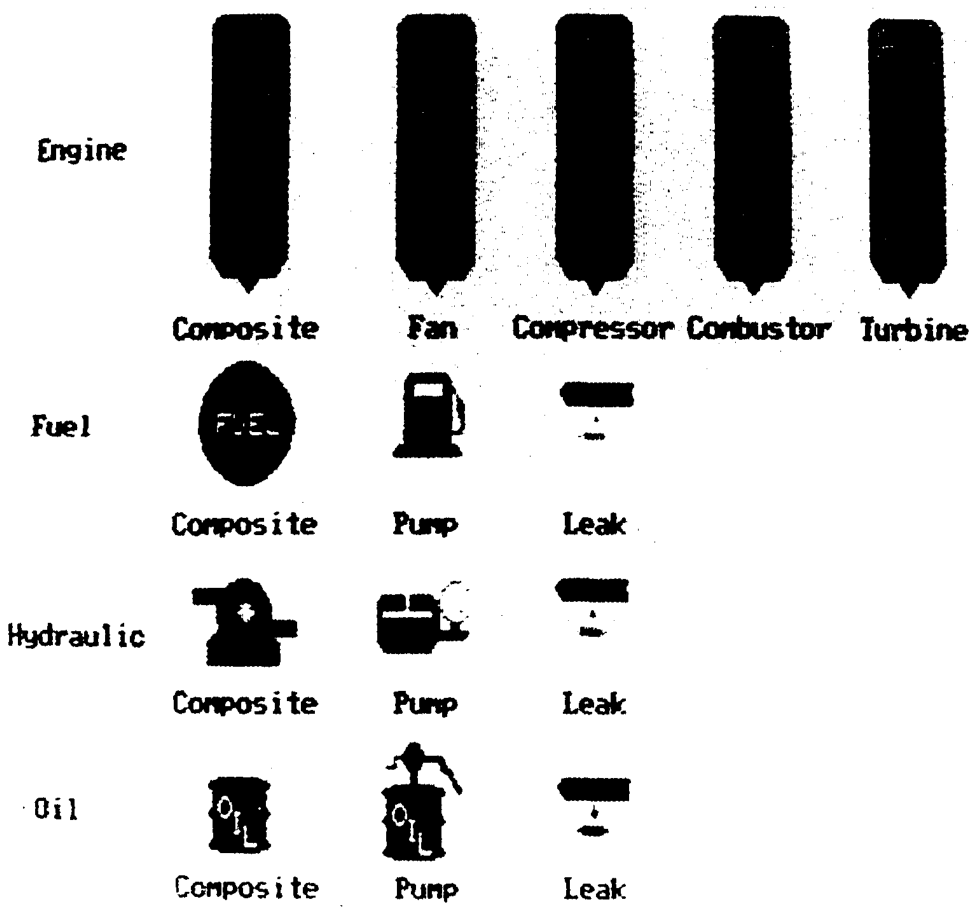
Symbols used in Displays



Appendix C

Pictures Used in Displays

Pictures Used in Displays



Appendix D

Written Instructions Given To Subjects

Appendix D

System Status Display Issues

A Study conducted by Don Allen in Cooperation with NASA-Langley Research Center

The present study is concerned with certain display issues that are related to the presentation of subsystem status information on CRT displays. The study is the first in a series of studies intended to investigate issues related to the introduction of advanced technology, principally expert system technology, into the commercial transport aircraft of the future. It is important to be aware of the preliminary nature of the displays you will see. You are one of the first people to see this material so it may be unfamiliar to you and may conflict with some of the current airline practices and FAA regulations. Please keep an open mind when participating in the study and work as best you can. Your performance as well as your comments will be used to refine this material for future studies.

The displays used in the study are based upon the output from Faultfinder, a fault diagnostic expert system for commercial aircraft under development at NASA-Langley Research Center. Faultfinder attempts to diagnose system faults during aircraft operation. The specifics of this process are not important to the present study, but if you are interested in further information it will be made available to you after the study. Faultfinder monitors system status information to detect faults and identify out-of-tolerance conditions. It does this by

(Appendix Continued)

generating hypotheses that account for the information that it has gathered. The issues addressed in this investigation focus on how best to display this information to the pilot.

The study will be conducted in three parts. Each part will consist of a set of displays based on a different style of information presentation. The three sets of displays are based on three different information presentation styles; text-, symbol-, or picture-based. The text-based displays take the form of an alphanumeric description of the current aircraft state. The symbol-based displays are based on geometric figures that represent the different aircraft subsystems and their components. The picture-based displays are composed of icons that represent the different subsystems. These will be described in greater detail later.

Since the set of displays are intended for future aircraft the sensor values used in the study are speculative. When viewing the displays it is more important for you to be aware of the qualitative condition of a subsystem parameter (abnormal/normal, low/high) then some specific sensor value. The displays are intended to represent a "snapshot" in the process of fault development. When considering the displays you should envision the aircraft as operating at a normal cruising altitude with all other aircraft systems operating normally.

Your task will be to determine which of the four aircraft subsystems in the study are represented in the display and then to identify them. The displays will be presented on the computer screen in front of you. In an actual cockpit these displays would be present in conjunction with

(Appendix Continued)

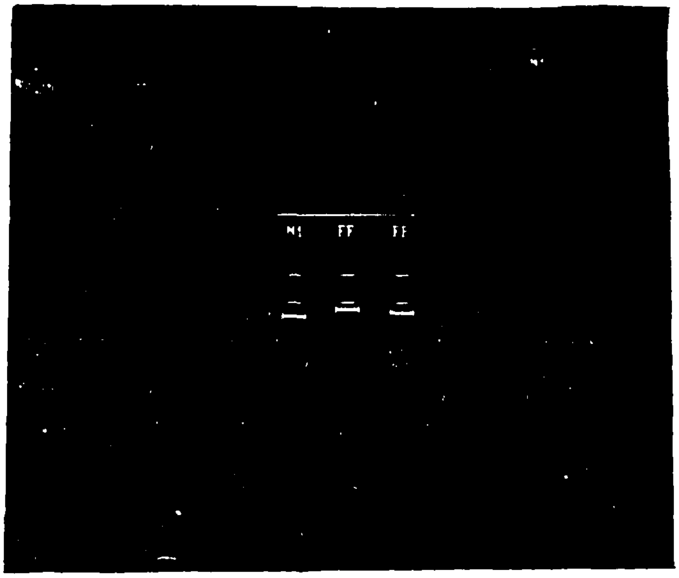
other subsystem display monitors. For this study you should only be concerned with the information presented in the generated displays. All other system status information should be considered normal. When the display appears on the screen you should determine the status of the aircraft. This should include the number and type of faulted subsystems displayed and the status of the different system parameters. When you have made these determinations you should depress the spacebar on the keyboard. This will cause the screen to go blank. You will then be asked to identify the different subsystems represented on the display and occasionally you will asked the operating status of a particular subsystem parameter. The time you need to determine the status of the aircraft will be monitored so it is important that you work as quickly as possible. However, do not sacrifice accuracy for speed. It is more important that you correctly identify the faulted subsystem(s) than it is for you to respond quickly. For some of the displays you will be asked a set of questions to determine how well the display conveyed system fault status. These questions will be asked randomly during the test sessions.

Each part of the experiment will consists of 64 displays (This may sound like a lot but you should move through them rapidly). For some of the displays, selected at random, you will be asked some questions about the display. These questions will concern current system state, future system state, and potential correcting actions. These questions are attempts to determine how well the system status information has been presented and not an attempt to determine your knowledge or piloting

(Appendix Continued)

skills. It is understood that the information that you might need you address the question in real flight situations would be contained in a flight manual. Please attempt to answer all of the questions to the best of your ability. When faced with a situation where you would require a flight manual to answer the question please indicate this to the experimenter. If you are not sure but can make an educated guess please feel free to do so, and indicate that it is a guess.

Each set of 64 displays will be presented in two blocks of 32 with a three minute break between blocks. One set of displays will take the following form:

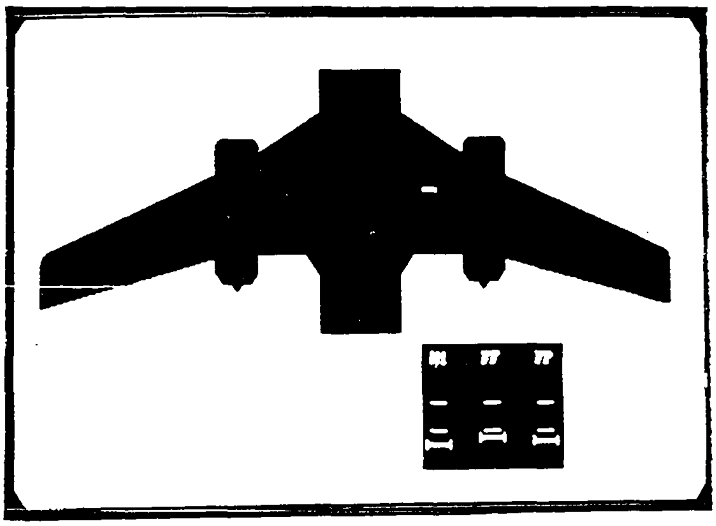


This display is a multiple hypotheses display. This display may contain from 1 to four hypotheses generated by faultfinder to explain current out-of-tolerance condition. Each of the hypothesis describes a fault or faults as they relate to the components of the aircraft subsystems used

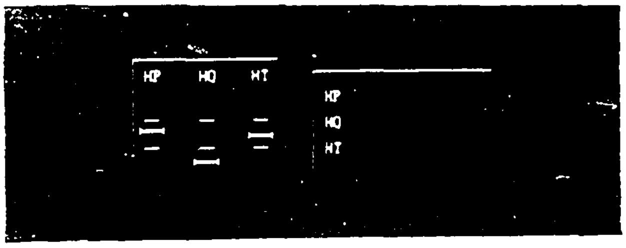
(Appendix Continued)

in the study (engine, oil, hydraulic, and fuel). Each block in the display describes one of the possible faults that explain the current condition of the aircraft (in this case a bad fuel pump). You will be asked to identify the subsystems represented by the components and not the individual faults themselves.

The second set of displays will be in the form of a composite hypothesis display. Rather than displaying individual components, these displays present the subsystems that are suspected of containing a fault. The elements of this set of displays represent subsystems and do not address individual subsystem components faults. An example of a composite hypothesis display is presented below:



All displays will contain some system parameter information. The location of this information will vary depending on the two types of displays described above. System parameter information will be presented either with quantitative information in the form of a numerical value or without this information. Examples of these are presented below:



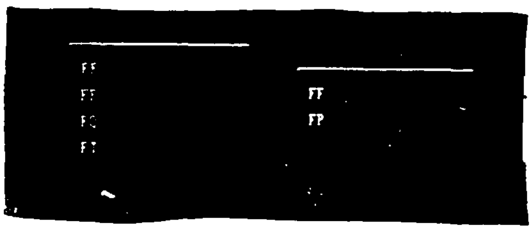
Without Quantitative Information	With Quantitative Information

The quantitative information presented in a red or green box depending on whether that value is considered normal or abnormal by Faultfinder. The green box denotes normal operation, while the red box denotes abnormal operation. When a value is considered abnormal a black arrowhead will appear next to the number to indicate whether it is abnormal-high or abnormal-low (in the display HQ is abnormal-low). For the qualitative displays the green area represents normal operation and

(Appendix Continued)

the two red regions represent abnormal high and abnormal low operation. The yellow regions denote caution regions between the two operating modes. The white bar represents the operation mode of the value.

The amount of parameter information will also differ. In some displays only abnormal values will be presented while for other displays all of the information related to the current fault will be displayed. The examples of these two situations are depicted below:



All Relevant Values   Abnormal Values

To recap. Your task in the current study is to determine the number and identity of subsystems involved in the out-of-tolerance condition identified by Faultfinder. This will involve viewing the display and pressing the spacebar when you understand what has been represented in the display. After you depress the spacebar you will be asked to identify the faulted subsystems depicted in the display. Some displays will be followed by an additional set of questions while others will not. To continue on to the next display during the study simply hit the ENTER key.

(Appendix Continued)

The types of displays you will see will vary. There are three major sets of displays: text-, symbol-, and picture-based. Within each set of displays their will be two blocks of 32 displays. One block will consist of displays presented in a multiple hypotheses form and the other block will consist of 32 displays presented in a composite hypothesis form. The blocks will be presented separately with a three minute break between them. The information presented in each block will differ in the type of parameter information (quantitative vs qualitative) and the amount of information (related information and abnormal information).

This study will only look at faults that pertain to four aircraft subsystems. These are:

Engine *: For our purposes the engine (referred to as the power plant in flight manuals) will consist of four components: fan, compressor, combustor, and the turbine. Each of these components could serve as a fault hypotheses in the multiple hypotheses displays. Five subsystem parameters are used in the study to describe the operation of the engine: N1 (speed of low pressure compressor), N2 (speed of the high pressure compressor), EPR (engine pressure ratio), EGT (engine gas temperature), and FF (fuel flow). Fuel flow will be considered both a part of the fuel subsystem and the engine since it plays a role in both subsystems.

Hydraulic Subsystem *: This subsystem consist of hydraulic pumps and hydraulic lines. Subsystem parameter information will be

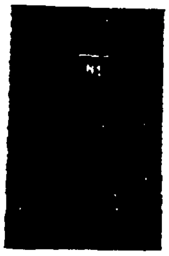
(Appendix Continued)

depicted as hydraulic pressure (HP), hydraulic quantity (HQ), and hydraulic temperature (HT).

Oils Subsystem *: Like the hydraulic subsystem the oil system will consist of oil pumps and lines. Subsystem parameter information will be represented as oil pressure (OP), oil quantity (OQ), and oil temperature (OT).

Fuel Subsystem *: This subsystem consists of fuel pumps and fuel lines. The fuel subsystem parameters are fuel flow (FF), fuel pressure (FP), fuel quantity (FQ), and fuel temperature (FT).

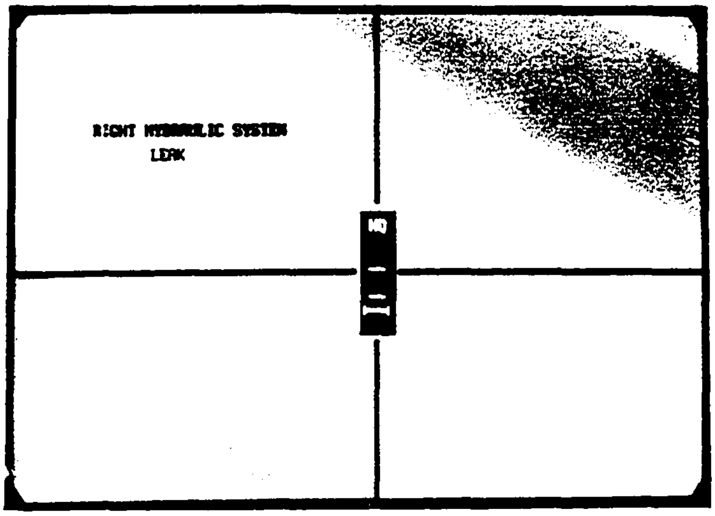
* The sensors that gather subsystem information are also subject to failure. Faultfinder is able to reason about potential faults occurring in the sensors. A faulty sensor is represented in the display below. Here a faulty N1 sensor is suspected.



(Appendix Continued)

Text-based Displays

This set of displays is based on the alphanumeric presentation of system fault information. These displays will be composed of brief verbal descriptions of the system state. This set of displays will be composed of multiple hypothesis situations that will look like the following:



Note that each of the individual hypotheses identifies a part of the subsystem that would be at fault for the current state of the subsystem. There may be some cases where a particular system state could be the result of a fault occurring in more than one subsystem. It is therefore important to be sure that you identify all subsystems in those instances.

To help you become familiar with the multiple hypotheses alphanumeric displays a number of example displays have been prepared for you to work through. You should work through the set in the same

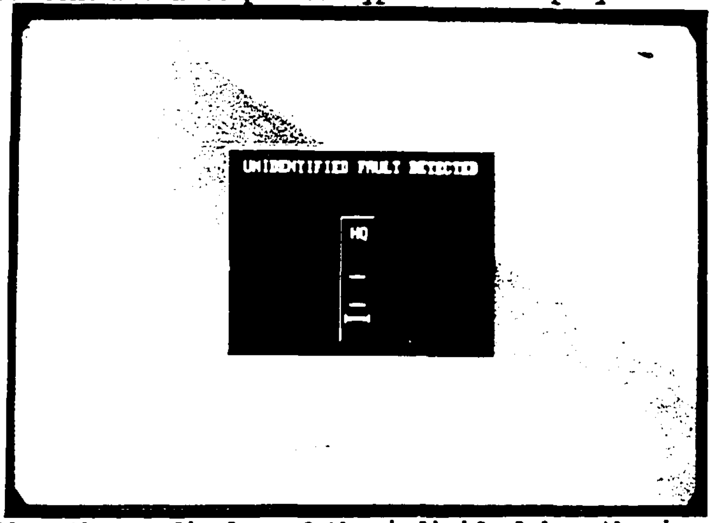
(Appendix Continued)

manner you would the actual displays. Once again, you need to look at the display and press the spacebar when you understand what has been represented in the display. The screen will then go blank and you be asked to identify the subsystems that have been represented on the screen. Some of the displays will be followed by a set of questions pertaining to that display. If you have any questions now or during this training session please asked the experimenter. If you are ready to proceed depress the ENTER key and begin.

(Appendix Continued)

Text-based Displays

The next set of displays are composed of composite hypothesis displays. A typical text-based composite hypothesis display will look like the following:



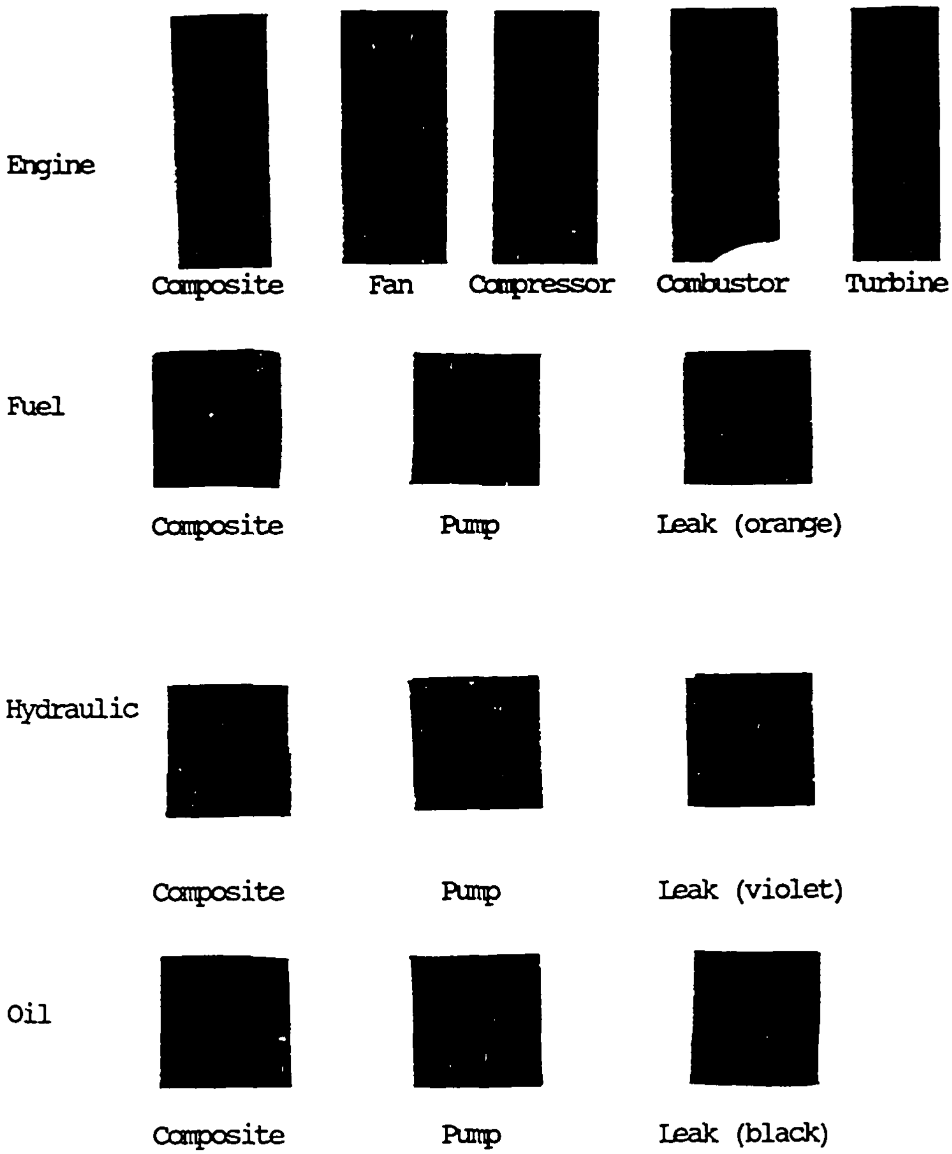
Notice that rather than a display of the individual hypothesis, the display contains the name of the subsystem(s) that is suspected of being at fault.

To help you become familiar with the set of composite hypotheses alphanumeric displays a number of example displays have been prepared for you to work through. You should work through the set in the same manner you would the actual displays used in the experiment. Once again, you need to look at the display and press the spacebar when you understand what has been represented in the display. The screen will then go blank and you be asked to identify the subsystems that have been represented on the screen. Some of the displays will be followed by a set of questions pertaining to that display. If you have any questions now or during this training session please asked the experimenter. If you are ready to proceed depress the ENTER key and begin.

(Appendix Continued)

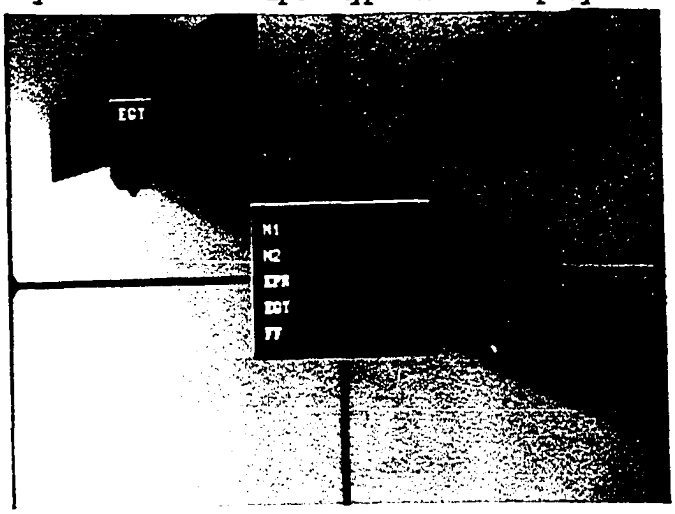
Symbol-based Displays

This set of displays is based on the presentation of system fault information as symbols. The displays will be composed of a set of geometric symbols that will be used to represent the four subsystems and their components. The subsystems and their associated components are presented below. Take some time to look them over.



(Appendix Continued)

A typical symbol-based multiple hypothesis display will look like the following:



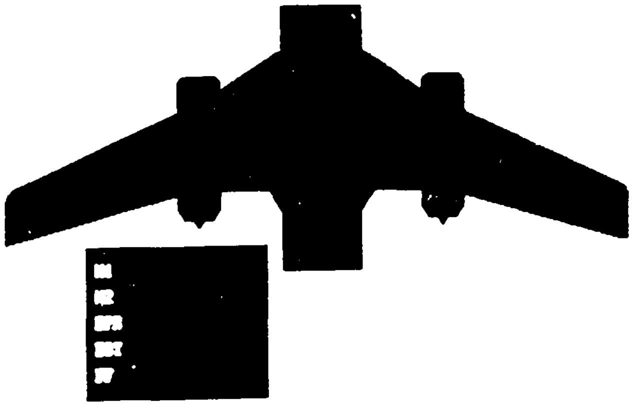
Note that each of the individual hypotheses identifies a part of the subsystem that could be at fault for the current state of the subsystem. There may be some cases where a particular system state could be the result of a fault in more than one subsystem. It is therefore important to be sure that you identify all subsystems in those instances.

To help you become familiar with this set of multiple hypotheses symbol displays a number of example displays have been prepared for you to work through. You should work through the set in the same manner you would the actual displays used in the experiment. Once again, you need to look at the display and press the spacebar when you understand what has been represented in the display. The screen will then go blank and you be asked to identify the subsystems that have been represented on the screen. Some of the displays will be followed by a set of questions pertaining to that display. If you have any questions now or during this training session please asked the experimenter. If you are ready to proceed depress the ENTER key.

(Appendix Continued)

Symbol-based Displays

The next set of displays are composed of composite hypothesis displays. Take time to look over the set of subsystem symbols again before continuing. A typical symbol-based composite hypothesis display will look like the following:



Notice that rather than a display of the individual hypotheses, the display contains the symbol of a subsystem(s) that is suspected of being at fault.

To help you become familiar with the set of composite hypothesis displays a number of example displays have been prepared for you to work through. You should work through the set in the same manner you would the actual displays used in the experiment. Once again, you need to look at the display and press the spacebar when you understand what has been represented in the display. The screen will then go blank and you be asked to identify the subsystems that have been represented on the screen. Some of the displays will be followed by a set of questions pertaining to that display. If you have any questions now or during this training session please asked the experimenter. If you are ready to proceed depress the ENTER key and begin.

(Appendix Continued)

Picture-based Displays

This set of displays is based on the presentation of system fault information as pictures or icons. These displays will be composed of a set of pictures that will be used to represent the four subsystems and their components. The subsystems and their components are presented below. Take some time to look them over.

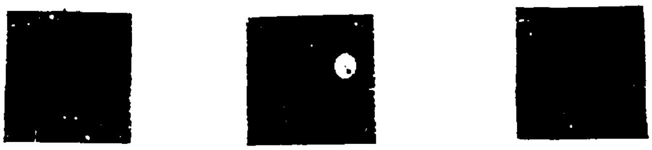
Engine

Composite Fan Compressor Combustor Turbine

Fuel

Composite Pump Leak (orange)

Hydraulic



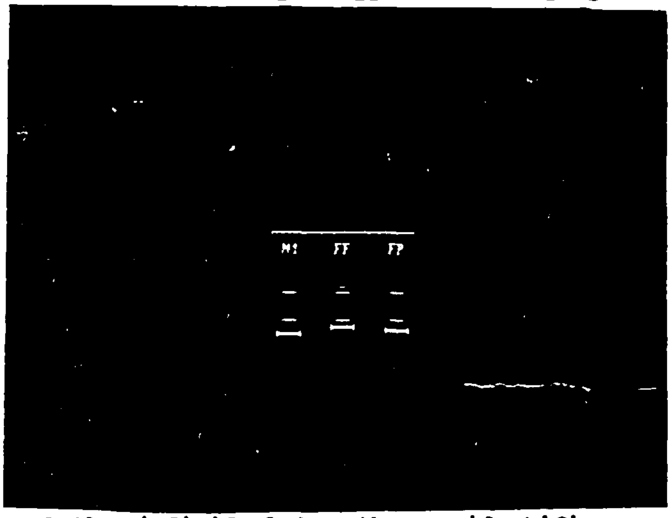
Composite Pump Leak (violet)

Oil

Composite Pump Leak (black)

(Appendix Continued)

A typical picture-based, multiple hypothesis display will look like the following:



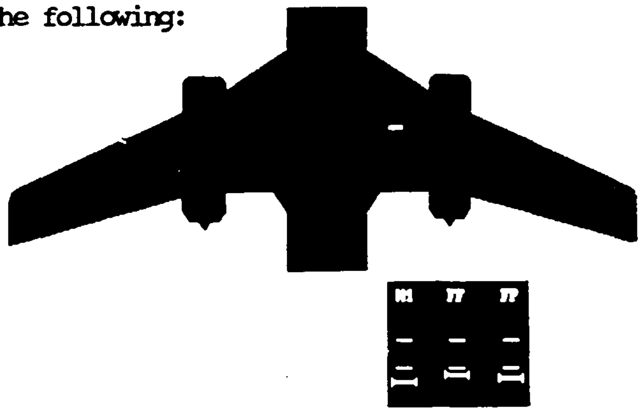
Note that each of the individual hypotheses identifies a part of the subsystem that could be at fault for the current state of the subsystem. There may be some cases where a particular system state could be the result of a fault in more than one subsystem. It is therefore important to be sure that you identify all subsystems in those instances.

To help you become familiar with this set of multiple hypotheses symbol displays a number of example displays have been prepared for you to work through. You should work through the set in the same manner you would the actual displays used in the experiment. Once again, you need to look at the display and press the spacebar when you understand what has been represented in the display. The screen will then go blank and you be asked to identify the subsystems that have been represented on the screen. Some of the displays will be followed by a set of questions pertaining to that display. If you have any questions now or during this training session please asked the experimenter. If you are ready to proceed depress the ENTER key an begin.

(Appendix Continued)

Picture-based Displays

The next set of displays will be composed of composite hypothesis displays. Take time to look over the set of subsystem picture elements again before continuing. A typical picture-based, composite hypothesis display will look like the following:



Notice that rather than a display of the individual hypotheses, the display contains the picture of a subsystem(s) that is suspected of being at fault.

To help you become familiar with the set of composite hypothesis displays a number of example displays have been prepared for you to work through. You should work through the set in the same manner you would the actual displays used in the experiment. Once again, you need to look at the display and press the spacebar when you understand what has been represented in the display. The screen will then go blank and you be asked to identify the subsystems that have been represented on the screen. Some of the displays will be followed by a set of questions pertaining to that display. If you have any questions now or during this training session please asked the experimenter. If you are ready to proceed depress the ENTER key and begin.

Appendix E

Post-Experimental Questionnaire Given To Subjects

Appendix E

Questionnaire

The following questions deal with the displays you have just seen. They are an attempt to get your preferences and to obtain any suggestions you may have concerning alternative display styles, other issues you feel need to be address, and your general comments.

1. Which of the three display types do you perfer? (1 - best)

   _____ Picture-base Displays

   _____ Symbol-based Displays

   _____ Text-based Displays

   Why? ____

   ____

   ____

   Why did you dislike the display you rated as "3"? ____________

   ____

   ____

   Can you suggest any alternative display styles? ____________

   ____

   ____

   Can you suggest any means for improving any of the three styles used in the study? ______________________________

   ____

   ____

(Appendix Continued)

2.Which type of parameter do you perfer?

_____ Qualitative Information (bargraphs)

_____ Quantitative Information (numerical values)

Why? ____

____

____

Why did you dislike the other type? ________________________

____

____

Can you suggest any alternative ways to display this information?

____

____

Comments: ____

____

____

3. Which type of hypotheses display type did you perfer?

_____ Composite Hypothesis

_____ Multiple Hypotheses

Why? ____

____

____

Why did you dislike the other type of display? ______________

____

____

(Appendix Continued)

Can you suggest any alternative ways to display the hypotheses?

____

____

Comments: ______________________________________

____

____

4. Did you perfer the presentation of all relevant data or just the abnormal data?

_____ Display of All Relevant Data

_____ Display of Relevant Out-of-Tolerance Data

Why? ____

____

____

Why did you dislike the presentation style? ______________

____

____

Are there other ways to segment this information? __________

____

____

Comments: ______________________________________

____

____

(Appendix Continued)

5. Were there any displays that you felt overwhelmed by? ____ If yes, please describe the display.____

____

____